PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
A01H 1/00, C07H 21/04, C07K 14/00, C12N 5/04, 5/10, C12P 19/34, C12Q 1/68

A1

(11) International Publication Number: WO 98/30083

(43) International Publication Date: 16 July 1998 (16.07.98)

(21) International Application Number: PCT/US98/00615

(22) International Filing Date: 9 January 1998 (09.01.98)

(30) Priority Data:
08/781,734 10 January 1997 (10.01.97) US

(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 22nd floor, 300 Lakeside Drive, Oakland, CA 94612 (US).

(72) Inventors: SHEN, Kathy; 44228 Country Club Drive, El Macero, CA 95618 (US). MEYERS, Blake; 904 Drake Drive, Davis, CA 95616 (US). MICHELMORE, Richard, W.; 36757 Russel Boulevard, Davis, CA 95616 (US).

(74) Agents: EINHORN, Gregory, P. et al.; Townsend and Townsend and Crew LLP, 8th floor, Two Embarcadero Center, San Francisco, CA 94111 (US).

(54) Title: RG NUCLEIC ACIDS FOR CONFERRING DISEASE RESISTANCE TO PLANTS

(57) Abstract

The present invention provides RG nucleic acids and proteins which confer disease resistance to plants. The nucleic acids can be used to produce transgenic plants resistant to pests. Antibodies to proteins of the invention are also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

RG NUCLEIC ACIDS FOR CONFERRING DISEASE RESISTANCE TO PLANTS

The present application is a continuation-in-part application ("CIP") of U.S. Patent Application Serial No. ("USSN") 08/781,734, filed January 10, 1997. The aforementioned application is explicitly incorporated herein by reference in its entirety and for all purposes.

This invention was made with Government support under Grant Nos. 92-37300-7547 and 95-37300-1571, awarded by the United States Department of Agriculture. The Government has certain rights in this invention.

FIELD OF THE INVENTION

The present invention relates generally to plant molecular biology. In particular, it relates to nucleic acids and methods for conferring pest resistance in plants. particularly lettuce.

BACKGROUND OF THE INVENTION

Recently, several resistance genes have been cloned by several groups from several plants. Many of these genes are sequence related. The derived amino acid sequences of the most common class, *RPS2, RPM1* (bacterial resistances in *Arabidopsis* (Mindrinos *et al. Cell* 78:1089-1099 (1994)); Bent *et al. Science* 265:1856-1860 (1994); Grant *et al.*, *Science* 269:843-846 (1995)), *L6* (fungal resistance in flax; Lawrence, *et al.*, *The Plant Cell* 7:1195-1206 (1995)), and *N*, (virus resistance in tobacco; Whitham, *et al.*, *Cell* 78:1101-1115 (1994); and U.S. Patent No. 5,571,706), all contain leucine-rich repeats (LRR) and nucleotide binding sites (NBS).

The NBS is a common motif in several mammalian gene families encoding signal transduction components (e.g., *Ras*) and is associated with ATP/GTP-binding sites.

The NBS is a common motif in several mammalian gene families encoding signal transduction components (e.g., *Ras*) and is associated with ATP/GTP-binding sites.

LRR domains can mediate protein-protein interactions and are found in a variety of proteins involved in signal transduction, cell adhesion and various other functions. LRRs are leucine rich regions often comprising 20-30 amino acid repeats where leucine and other aliphatic residues occur periodically. LRRs can function extracellularly or intracellularly.

Since the onset of civilization, plant diseases have had catastrophic effects on crops and the well-being of the human population. Plant diseases continue to effect enormous human and economic costs. An increasing human population and decreasing amounts of arable land make all approaches to preventing and treating plant pathogen destruction critical. The ability to control and enhance a plant's protective responses against pathogens would be of enormous benefit. Tissue-specific and temporal control of mechanisms responsible for plant cell death would also be of great practical and economic value. The present invention fulfills these and other needs.

What is needed in the art are plant disease resistance genes and means to create transgenic disease resistance plants, particularly in lettuce. Further, what is needed in the art is a means to DNA fingerprint cultivars and germplasm with respect to their disease resistance haplotypes for use in plant breeding programs. The present invention provides these and other advantages.

SUMMARY OF THE INVENTION

The present invention provides isolated nucleic acid constructs. These constructs comprise an RG (resistance gene) polynucleotide which encodes an RG polypeptide having at least 60% sequence identity to an RG polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, and an RG4 polypeptide. RG1, RG2, RG3, RG4, and the like, represent individual "RG families." Each "RG family," as defined herein, is a group of polypeptide sequences that have at least 60% amino acid sequence identity. Individual members of an RG family, *i.e.*, individual species of the genus, typically map to the same genomic locus. The invention provides for constructs comprising nucleotides encoding the RG families of the

invention, which can include sequences encoding a leucine rich region (LRR), and/or a nucleotide binding site (NBS), or both.

The invention provides for an isolated nucleic acid construct comprising an RG polynucleotide which encodes an RG polypeptide having at least 60% sequence identity to an RG polypeptide from an RG family selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide. In alternative embodiments, the nucleic acid construct comprises an RG polynucleotide which encodes an RG polypeptide comprising an leucine rich region (LRR), or, an RG polypeptide comprising a nucleotide binding site (NBS). The nucleic acid construct can comprise a polynucleotide which is a full length gene. In another embodiment, the nucleic acid construct encodes a fusion protein.

In one embodiment, the nucleic acid construct comprises a sequence encoding an RG1 polypeptide. The RG1 polypeptide can be encoded by a polynucleotide sequence selected from the group consisting of SEQ ID NO:1 (RG1A), SEQ ID NO:2 and SEQ ID NO:137 (RG1B), SEQ ID NO: 3 (RG1C), SEQ ID NO:4 (RG1D), SEQ ID NO:5 (RG1E), SEQ ID NO:6 (RG1F), SEQ ID NO:7 (RG1G), SEQ ID NO:8 (RG1H), SEQ ID NO:9 (RG1I), and SEQ ID NO:10 (RG1J).

In another embodiment, the nucleic acid construct comprises a sequence encoding an RG2 polypeptide. The RG2 polypeptide can be encoded by a polynucleotide sequence selected from the group consisting of: SEQ ID NO:21 and SEQ ID NO:27 (RG2A); SEQ ID NO:23 and SEQ ID NO:28 (RG2B); SEQ ID NO:29 (RG2C); SEQ ID NO:30 (RG2D); SEQ ID NO:31 (RG2E); SEQ ID NO:32 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:34 (RG2H); SEQ ID NO:35 (RG2I); SEQ ID NO:36 (RG2J); SEQ ID NO:37 (RG2K); SEQ ID NO:38 (RG2L); SEQ ID NO:39 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).

In other embodiments, the nucleic acid construct comprises a RG3 sequence (SEQ ID NO:68) encoding an RG3 polypeptide (SEQ ID NO:138) (RG3). In other embodiments, the nucleic acid construct comprises an RG4 sequence (SEQ ID NO:69) encoding an RG4 polypeptide (SEQ ID NO:139) (RG4).

In other embodiments, the nucleic acid construct comprises a RG5 sequence (SEQ ID NO:134) encoding an RG5 polypeptide (SEQ ID NO:135). The RG5 polypeptide can be encoded by a polynucleotide sequence as set forth in SEQ ID NO:134.

The invention also provides for a nucleic acid construct which comprises an RG7 sequence encoding an RG7 polypeptide. The RG7 polypeptide can be encoded by a polynucleotide sequence as set forth in SEQ ID NO:136.

In further embodiments, the nucleic acid construct can further comprise a promoter operably linked to the RG polynucleotide. In alternative embodiments, the promoter can be a plant promoter; a disease resistance promoter; a lettuce promoter; a constitutive promoter; an inducible promoter; or, a tissue-specific promoter. The nucleic acid construct can comprise a promoter sequence from an RG gene linked to a heterologous polynucleotide.

The invention also provides for a transgenic plant comprising a recombinant expression cassette comprising a promoter operably linked to an RG polynucleotide. The expression cassette can comprise a plant promoter or a viral promoter; the plant promoter can be a heterologous promoter. In one embodiment, the transgenic plant is lettuce. In alternative embodiments, the transgenic plant comprises an expression cassette which includes an RG polynucleotide selected from the group consisting of SEQ ID NO:1 (RG1A); SEQ ID NO:2 and SEQ ID NO:137 (RG1B); SEQ ID NO: 3 (RG1C); SEQ ID NO:4 (RG1D); SEQ ID NO:5 (RG1E); SEQ ID NO:6 (RG1F); SEQ ID NO:7 (RG1G); SEQ ID NO:8 (RG1H); SEQ ID NO:9 (RG1I) and SEQ ID NO:10 (RG1J); SEQ ID NO:21 and SEQ ID NO:27 (RG2A); SEQ ID NO:23 and SEQ ID NO:28 (RG2B); SEQ ID NO:29 (RG2C); SEQ ID NO:30 (RG2D); SEQ ID NO:31 (RG2E); SEQ ID NO:32 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:34 (RG2H); SEQ ID NO:35 (RG2I); SEQ ID NO:36 (RG2J); SEQ ID NO:37 (RG2K); SEQ ID NO:38 (RG2L); SEQ ID NO:39 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104

(RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W); SEQ ID NO:68 (RG3); SEQ ID NO:69 (RG4); SEQ ID NO:134 (RG5); or SEQ ID NO:136 (RG7).

The invention provide for a transgenic plant comprising an expression cassette comprising an RG polynucleotide which can encode an RG1 polypeptide selected from the group consisting of SEQ ID NO:11 (RG1A), SEQ ID NO:12 (RG1B), SEQ ID NO:13 (RG1C), SEQ ID NO:14 (RG1D), SEQ ID NO:15 (RG1E), SEQ ID NO:16 (RG1F), SEQ ID NO:17 (RG1G), SEQ ID NO:18 (RG1H), SEQ ID NO:19 (RG1I), or SEQ ID NO:20 (RG1J); or, an RG2 polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W); an RG4 polypeptide as set forth by SEQ ID NO:72; an RG5 polypeptide with a sequence as set forth by SEQ ID NO:135; or, an RG7 polypeptide.

The invention also provides for a method of enhancing disease resistance in a plant, the method comprising introducing into the plant a recombinant expression cassette comprising a promoter functional in the plant and operably linked to an RG polynucleotide sequence. In this method, the plant can be a lettuce plant; and, the RG polynucleotide can encode an RG polypeptide selected from the group consisting of an RG1 polypeptide selected from the group consisting of SEQ ID NO:11 (RG1A), SEQ ID NO:12 (RG1B), SEQ ID NO:13 (RG1C), SEQ ID NO:14 (RG1D), SEQ ID NO:15 (RG1E), SEQ ID

NO:16 (RG1F), SEQ ID NO:17 (RG1G), SEQ ID NO:18 (RG1H), SEQ ID NO:19 (RG1I), or SEQ ID NO:20 (RG1J); or, an RG2 polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:72; SEQ ID NO:74; SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W). In this method, the promoter can be a plant disease resistance promoter, a tissue-specific promoter, a constitutive promoter, or an inducible promoter.

The invention also provides for a method of detecting RG resistance genes in a nucleic acid sample, the method comprising: contacting the nucleic acid sample with an RG polynucleotide to form a hybridization complex; and, wherein the formation of the hybridization complex is used to detect the RG resistance gene in the nucleic acid sample. In this method, the RG polynucleotide can be an RG1 polynucleotide, an RG2 polynucleotide, an RG3 polynucleotide, an RG4 polynucleotide, an RG5 polynucleotide or an RG7 polynucleotide. In this method, the RG resistance gene can be amplified prior to the step of contacting the nucleic acid sample with the RG polynucleotide, and, the RG resistance gene can be amplified by the polymerase chain reaction. In one embodiment, the RG polynucleotide is labeled.

The invention further provides for an RG polypeptide having at least 60% sequence identity to a polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide.

A further understanding of the nature and advantages of the present invention may be realized by reference to the remaining portions of the specification, the figures and claims.

All publications, patents and patent applications cited herein are hereby expressly incorporated by reference for all purposes.

DETAILED DESCRIPTION OF THE INVENTION

This invention relates to families of RG genes, particularly from *Lactuca sativa*. Nucleic acid sequences of the present invention can be used to confer resistance in plants to a variety of pests including viruses, fungi, nematodes, insects, and bacteria. Sequences from within the RG genes can be used to fingerprint cultivars or germplasm for the presence of desired resistance genes. Promoters of RG genes can be used to drive heterologous gene expression under conditions in which RG genes are expressed. Further, the present invention provides RG proteins and antibodies specifically reactive to RG proteins. Antibodies to RG proteins can be used to detect the type and amount of RG protein expressed in a plant sample.

The present invention has use over a broad range of types of plants, including species from the genera *Cucurbita, Rosa, Vitis, Juglans, Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon, Nicotiana, Solanum, Petunia, Digitalis, Majorana, Ciahorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Heterocallis, Nemesis, Pelargonium, Panieum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browaalia, Glycine, Pisum, Phaseolus, Lolium, Oryza, Zea, Avena, Hordeum, Secale, Triticum*, and, *Sorghum*. In particularly preferred embodiments, species from the family *Compositae* and in particular the genus *Lactuca* are employed such as *L. sativa* and such subspecies as *crispa, longifolia,* and *asparagina*.

The nucleic acids of the present invention can be used in marker-aided selection. Marker-aided selection does not require the complete sequence of the gene or precise knowledge of which sequence confers which specificity. Instead, partial sequences can be used as hybridization probes or as the basis for oligonucleotide primers to amplify nucleic acid, e.g., by PCR. Partial sequences can be used in other methods, such as to

follow the segregation of chromosome segments containing resistance genes in plants. Because the RG marker is the gene itself, there can be negligible recombination between the marker and the resistance phenotype. Thus, RG polynucleotides of the present invention provide an optimal means to DNA fingerprint cultivars and wild germplasm with respect to their disease resistance haplotypes. This can be used to indicate which germplasm accessions and cultivars carry the same resistance genes. At present, selection of plants (e.g., lettuce) for resistance to some diseases is slow and difficult. But linked markers allow indirect selection for such resistance genes. Moreover, RG markers also allow resistance genes to be identified and combined in a manner that would not otherwise be possible. Numerous accessions have been identified that provide resistance to all isolates of downy mildew (*Bremia lactucae*). However, without molecular markers it is impossible to combine such resistances from different sources. The nucleic acid sequences of the invention provide for a fast and convenient means to identify and combine resistances from different sources. The RG markers of the invention can also be used to identify recombinants that have new combinations of resistance genes in *cis* on the same chromosome.

In addition, RG markers may allow the identification of the Mendelian factors determining traits, such as field resistance to downy mildew. Once such markers have been identified, they will greatly increase the ease with which field resistance can be transferred between lines and combined with other resistances.

In another application, primers to RG sequences can be also designed to amplify sequences that are conserved in multiple RG family members. This gives genetic information on multiple RG family members. Alternatively, one or more primers can be made to sequences unique to a single resistance gene genus or a single RG specie. This allows an analysis of individual family groups (an RG genus) or an individual family member (a specie). Primers made to individual RGs at the edge of each cluster can be used to select for recombinants within the cluster. This minimizes the amount of linkage drag during introgression. Classical and molecular genetics has shown that pest resistance genes tend to be clustered in the genome. Pest resistance loci comprise arrays of genes and exhibit a variety of complex haplotypes rather than being simple alternate allelic forms. Pest resistance is conferred by families, or genuses, of related RG sequences, individual members, or species, of which have evolved to have a different specificity.

Oligonucleotide primers can be designed that amplify members from multiple haplotypes, or genuses, or amplify only members of one genus, or only amplify an individual specie. This will provide codominant information and allow heterozygotes to be distinguished from homozygotes.

Further, comparison of RG sequences will allow a determination of which sequences are critical for resistance and will ultimately lead to engineering resistance genes with new specificities. Resistance gene sequences were not previously available for lettuce. Marker-aided selection will greatly increase the precision and speed of breeding for disease resistance. Transgenic approaches will allow pyramiding of resistance genes into a single Mendelian unit, transfer between sexually-incompatible species, substitute for conventional backcrossing procedures, and allow expression of other genes in parallel with resistance genes.

The RG polynucleotides also have utility in the construction of disease resistant transgenic plants. This avoids lengthy and sometimes difficult backcrossing programs currently necessary for introgression of resistance. It is also possible to transfer resistance polynucleotides between sexually-incompatible species, thereby greatly increasing the germplasm pool that can be used as a source of resistance genes. Cloning of multiple RG sequences in a single cassette will allow pyramiding of genes for resistance against multiple isolates of a single pathogen such as downy mildew or against multiple pathogens. Once introduced, such a cassette can be manipulated by classical breeding methods as a single Mendelian unit.

Transgenic plants of the present invention can also be constructed using an RG promoter. The promoter sequences from RG sequences of the invention can be used with RG genes or heterologous genes. Thus, RG promoters can be used to express a variety of genes in the same temporal and spatial patterns and at similar levels to resistance genes.

Nucleic acids of the Invention and Their Preparation

RG Polynucleotide Families

The present invention provides isolated nucleic acid constructs which comprise an RG polynucleotide. In alternative embodiments, the RG polynucleotide is at least 18 nucleotides in length, typically at least 20, 25, or 30 nucleotides in length, more

typically at least 100 nucleotides in length, generally at least 200 nucleotides in length, preferably at least 300 nucleotides in length, more preferably at least 400 nucleotides in length, and most preferably at least 500 nucleotides in length.

In particularly preferred embodiments, the RG polynucleotide encodes a RG protein which confers resistance to plant pests. This RG protein can be longer, equivalent, or shorter than the RG protein encoded by an RG gene. In various embodiments, an RG polynucleotide can hybridize under stringent conditions to members of an RG family (an RG genus); *e.g.*, it can hybridize to a member of the RG1 RG family, such as an RG1 polynucleotide selected from the group consisting of: SEQ ID NO:1 (RG1A); SEQ ID NO:2 and SEQ ID NO:137 (RG1B); SEQ ID NO: 3 (RG1C); SEQ ID NO:4 (RG1D); SEQ ID NO:5 (RG1E); SEQ ID NO:6 (RG1F); SEQ ID NO:7 (RG1G); SEQ ID NO:8 (RG1H); SEQ ID NO:9 (RG1I) and SEQ ID NO:10 (RG1J).

In other embodiments, the polynucleotide can also hybridize under stringent conditions to a member of the RG2 family; such as an RG2 polynucleotide selected from the group consisting of: SEQ ID NO:21 and SEQ ID NO:27 (RG2A); SEQ ID NO:23 and SEQ ID NO:28 (RG2B); SEQ ID NO:29 (RG2C); SEQ ID NO:30 (RG2D); SEQ ID NO:31 (RG2E); SEQ ID NO:32 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:34 (RG2H); SEQ ID NO:35 (RG2I); SEQ ID NO:36 (RG2J); SEQ ID NO:37 (RG2K); SEQ ID NO:38 (RG2L); SEQ ID NO:39 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).

In alternative embodiments, each RG2 gene can also include an AC15 sequence which hybridizes under stringent conditions to a polynucleotide selected from the group consisting of: SEQ ID NO:56 (AC15-2A); SEQ ID NO:57 (AC15-2B); SEQ ID NO:58 (AC15-2C); SEQ ID NO:59 (AC15-2D); SEQ ID NO:60 (AC15-2E); SEQ ID NO:61 (AC15-2G); SEQ ID NO:62 (AC15-2H); SEQ ID NO:63 (AC15-2I); SEQ ID

NO:64 (AC15-2J); SEQ ID NO:65 (AC15-2L); SEQ ID NO:66 (AC15-2N); SEQ ID NO:67 (AC15-2O).

In other embodiments, an RG polynucleotide can hybridize under stringent conditions to an RG3 (SEQ ID NO:68), an RG4 (SEQ ID NO:69), and RG5 (SEQ ID NO:135), and an RG7 (SEQ ID NO:137), RG family member.

The present invention further provides nucleic acid constructs which comprise an RG polynucleotide which encodes RG polypeptides from various RG families; such as an RG polypeptide having at least 60% sequence identity to an RG polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, and RG4 polypeptide, and RG5 polypeptide, and an RG7 polypeptide.

Exemplary RG1 polypeptides have the sequences shown in SEQ ID NO:2 (RG1A), SEQ ID NO:4 (RG1B), SEQ ID NO:6 (RG1C), SEQ ID NO:8 (RG1D), SEQ ID NO:10 (RG1E), SEQ ID NO:12 (RG1F), SEQ ID NO:14 (RG1G), SEQ ID NO:16 (RG1H), SEQ ID NO:20 (RG1J). Exemplary RG2 polypeptides have the sequences shown in SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W).

An exemplary RG3 polypeptide has the sequence shown in SEQ ID NO:138. An exemplary RG4 polypeptide has the sequence shown in SEQ ID NO:139. RG polynucleotides will have at least 60% identity, more typically at least 65% identity, generally at least 70% identity, and preferably at least 75% identity, more preferably at least 80% identity, and most preferably at least 85%, 90%, or 95% identity at the deduced amino acid level. The regions where substantial identity is assessed can be inclusive or exclusive of the nucleotide binding site or the leucine rich region.

Vectors and Transcriptional Control Elements

The invention, providing methods and reagents for making novel species and genuses of RG nucleic acids described herein, further provides methods and reagents for expressing these nucleic acids using novel expression cassettes, vectors, transgenic plants and animals, using constitutive and inducible transcriptional and translational *cis*- (*e.g.*. promoters and enhancers) and *trans*-acting control elements.

The expression of natural, recombinant or synthetic plant disease resistance polypeptide-encoding or other (*i.e.*, antisense, ribozyme) nucleic acids can be achieved by operably linking the coding region a promoter (that can be plant-specific or not, constitutive or inducible), incorporating the construct into an expression cassette (such as an expression vector), and introducing the resultant construct into an *in vitro* reaction system or a suitable host cell or organism. Synthetic procedures may also be used. Typical expression systems contain, in addition to coding or antisense sequence, transcription and translation terminators, polyadenylation sequences, transcription and translation initiation sequences, and promoters useful for transcribing DNA into RNA. The expression systems optionally at least one independent terminator sequence, sequences permitting replication of the cassette *in vivo, e.g.*, plants, eukaryotes, or prokaryotes, or a combination thereof, (*e.g.*, shuttle vectors) and selection markers for the selected expression system, *e.g.*, plant, prokaryotic or eukaryotic systems. To ensure proper polypeptide expression under varying conditions, a polyadenylation region at the 3'-end of the coding region can be included (see Li (1997) *Plant Physiol.*115:321-325, for a review of the polyadenylation of RNA in plants). The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA (*e.g.*, using *Agrobacterium tumefaciens* T-DNA replacement vectors, see *e.g.*, Thykjaer (1997) *Plant Mol Biol.* 35:523-530; using a plasmid containing a gene of interest flanked by *Agrobacterium* T-DNA border repeat sequences; Hansen (1997) "T-strand integration in maize protoplasts after codelivery of a T-DNA substrate and virulence genes," *Proc. Natl. Acad. Sci. USA* 94:11726-11730.

To identify the promoters, the 5' portions of the clones described here are analyzed for sequences characteristic of promoter sequences. For instance, promoter sequence elements include the TATA box consensus sequence (TATAAT), which is usually 20 to 30 base pairs upstream of the transcription start site. In plants, further

upstream from the TATA box, at positions -80 to -100, there is typically a promoter element with a series of adenines surrounding the trinucleotide G (or T) N G (see, *e.g.*, Messing, in *Genetic Engineering in Plants*, pp. 221-227, Kosage, Meredith and Hollaender, eds. 1983). If proper polypeptide expression is desired, a polyadenylation region at the 3'-end of the RG coding region should be included. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from viral genes, such as T-DNA.

The nucleic acids of the invention can be expressed in expression cassettes, vectors or viruses which are transiently expressed in cells using, for example, episomal expression systems (*e.g.*, cauliflower mosaic virus (CaMV) viral RNA is generated in the nucleus by transcription of an episomal minichromosome containing supercoiled DNA, Covey (1990) *Proc. Natl. Acad. Sci. USA* 87:1633-1637). Alternatively, coding sequences can be inserted into the host cell genome becoming an integral part of the host chromosomal DNA.

Selection markers can be incorporated into expression cassettes and vectors to confer a selectable phenotype on transformed cells and sequences coding for episomal maintenance and replication such that integration into the host genome is not required. For example, the marker may encode biocide resistance, such as antibiotic resistance, particularly resistance to chloramphenicol, kanamycin, G418, bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosulfuron or Basta, to permit selection of those cells transformed with the desired DNA sequences, see for example, Blondelet-Rouault (1997) *Gene* 190:315-317; Aubrecht (1997) *J. Pharmacol. Exp. Ther.* 281:992-997 . Because selectable marker genes conferring resistance to substrates like neomycin or hygromycin can only be utilized in tissue culture, chemoresistance genes are also used as selectable markers *in vitro* and *in vivo*. See also, Mengiste (1997) "High-efficiency transformation of *Arabidopsis thaliana* with a selectable marker gene regulated by the T-DNA 1' promoter," *Plant J.* 12:945-948, showing that the 1' promoter is an attractive alternative to the cauliflower mosaic virus (CaMV) 35S promoter for the generation of T-DNA insertion lines, the 1' promoter may be especially beneficial for the secondary transformation of transgenic strains containing the 35S promoter to exclude homology-mediated gene silencing.

The endogenous promoters from the RG genes of the present invention can be used to direct expression of the genes. These promoters can also be used to direct expression of heterologous structural genes. The promoters can be used, for example, in recombinant expression cassettes to drive expression of genes conferring resistance to any number of pathogens or pests, including fungi, bacteria, and the like.

Constitutive Promoters

In construction of recombinant expression cassettes, vectors, transgenics, of the invention, a promoter fragment can be employed to direct expression of the desired gene in all tissues of a plant or animal. Promoters that drive expression continuously under physiological conditions are referred to as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include those from viruses which infect plants, such as the cauliflower mosaic virus (CaMV) 35S transcription initiation region; the 1'- or 2'-promoter derived from T-DNA of *Agrobacterium tumafaciens*; the promoter of the tobacco mosaic virus; and, other transcription initiation regions from various plant genes known to those of skill. See also Holtorf (1995) "Comparison of different constitutive and inducible promoters for the overexpression of transgenes in *Arabidopsis thaliana*," *Plant Mol. Biol.* 29:637-646.

Inducible Promoters

Alternatively, a plant promoter may direct expression of the plant disease resistance nucleic acid of the invention under the influence of changing environmental conditions or developmental conditions. Examples of environmental conditions that may effect transcription by inducible promoters include pathogenic attack, anaerobic conditions, elevated temperature, drought, or the presence of light. Such promoters are referred to herein as "inducible" promoters. For example, the invention incorporates the drought-inducible promoter of maize (Busk (1997) *supra*); the cold, drought, and high salt inducible promoter from potato (Kirch (1997) *Plant Mol. Biol.* 33:897-909).

Embodiments of the invention also incorporate use of plant promoters which are inducible upon injury or infection to express the invention's plant disease resistance (RG) polypeptides. Various embodiments include use of, *e.g.*, the promoter for a tobacco (*Nicotiana tabacum*) sesquiterpene cyclase gene (EAS4 promoter), which is expressed in wounded leafs, roots, and stem tissues, and upon infection with microbial pathogens (Yin

(1997) *Plant Physiol.* 115(2):437-451); the ORF13 promoter from *Agrobacterium rhizogenes* 8196, which is wound inducible in a limited area adjacent to the wound site (Hansen (1997) *Mol. Gen. Genet.* 254:337-343); the Shpx6b gene promoter, which is a plant peroxidase gene promoter induced by microbial pathogens (demonstrated using a fungal pathogen, see Curtis (1997) *Mol. Plant Microbe Interact.* 10:326-338); the wound-inducible gene promoter wun1, derived from potato (Siebertz (1989) *Plant Cell* 1:961-968); the wound-inducible *Agrobacterium pmas* gene (mannopine synthesis gene) promoter (Guevara-Garcia (1993) *Plant J.* 4:495-505).

Alternatively, plant promoters which are inducible upon exposure to plant hormones, such as auxins, are used to express the nucleic acids of the invention. For example, the invention can use the auxin-response elements E1 promoter fragment (AuxREs) in the soybean (*Glycine max L.*) (Liu (1997) Plant Physiol. 115:397-407); the auxin-responsive Arabidopsis GST6 promoter (also responsive to salicylic acid and hydrogen peroxide) (Chen (1996) *Plant J.* 10: 955-966); the auxin-inducible parC promoter from tobacco (Sakai (1996) 37:906-913); a plant biotin response element (Streit (1997) *Mol. Plant Microbe Interact.* 10:933-937); and, the promoter responsive to the stress hormone abscisic acid (Sheen (1996) *Science* 274:1900-1902).

Plant promoters which are inducible upon exposure to chemicals reagents which can be applied to the plant, such as herbicides or antibiotics, are also used to express the nucleic acids of the invention. For example, the maize In2-2 promoter, activated by benzenesulfonamide herbicide safeners, can be used (De Veylder (1997) Plant Cell Physiol. 38:568-577); application of different herbicide safeners induces distinct gene expression patterns, including expression in the root, hydathodes, and the shoot apical meristem. Coding sequence can be under the control of, *e.g.*, a tetracycline-inducible promoter, *e.g.*, as described with transgenic tobacco plants containing the *Avena sativa* L. (oat) arginine decarboxylase gene (Masgrau (1997) *Plant J.* 11:465-473); or, a salicylic acid-responsive element (Stange (1997) *Plant J.* 11:1315-1324. Using chemically- (*e.g.*, hormone- or pesticide-) induced promoters, harvesting of fruits and plant parts would be greatly facilitated. A chemical which can be applied to the transgenic plant in the field and induce expression of a polypeptide of the invention throughout all or most of the plant would make a environmentally safe defoliant or herbicide. Thus, the invention also provides for transgenic plants containing an inducible gene encoding for the RG

polypeptides of the invention whose host range is limited to target plant species, such as weeds or crops before, during or after harvesting.

Abcission promoters are activated upon plant ripening, such as fruit ripening, and are especially useful incorporated in the expression systems (*e.g.*, expression cassettes, vectors) of the invention. In some embodiments, when a plant disease resistant polypeptide-encoding nucleic acid is under the control of such a promoter, rapid cell death, induced by expression of the invention's polypeptide, can accelerate and/or accentuate abcission, increasing the efficiency of the harvesting of fruits or other plant parts, such as cotton, and the like. Induction of rapid cell death at this time would accelerate separation of the fruit from the plant, greatly augmenting harvesting procedures. See, *e.g.*, Kalaitzis (1997) *Plant Physiol.* 113:1303-1308, discussing tomato leaf and flower abscission; Payton (1996) *Plant Mol. Biol.* 31:1227-1231, discussing ethylene receptor expression regulation during fruit ripening, flower senescence and abscission; Koehler (1996) *Plant Mol. Biol.* 31:595-606, discussing the gene promoter for a bean abscission cellulase; Kalaitzis (1995) *Plant Mol. Biol.*28: 647-656, discussing cloning of a tomato polygalacturonase expressed in abscission; del Campillo (1996) *Plant Physiol.* 111:813-820, discussing pedicel breakstrength and cellulase gene expression during tomato flower abscission.

Tissue-Specific Promoters

Tissue specific promoters are transcriptional control elements that are only active in particular cells or tissues. Plant promoters which are active only in specific tissues or at specific times during plant development are used to express the nucleic acids of the invention. Examples of promoters under developmental control include promoters that initiate transcription only in certain tissues, such as leaves, roots, fruit, seeds, ovules, pollen, pistols, or flowers. Such promoters are referred to as "tissue specific". The operation of a promoter may also vary depending on its location in the genome. Thus, an inducible promoter may become fully or partially constitutive in certain locations.

For example, a seed-specific promoter directs expression in seed tissues. Such promoters may be, for example, ovule-specific, embryo-specific, endosperm-specific, integument-specific, seed coat-specific, or some combination thereof. A leaf-specific promoter has been identified in maize, Busk (1997) *Plant J.* 11:1285-1295. The ORF13 promoter from *Agrobacterium rhizogenes* exhibits high activity in roots (Hansen (1997) *supra*). A maize pollen-specific promoter has been identified in maize (Guerrero (1990)

Mol. Gen. Genet. 224:161-168). A tomato promoter active during fruit ripening, senescence and abscission of leaves and, to a lesser extent, of flowers can be used (Blume (1997) *Plant J.* 12:731-746). A pistol specific promoter has been identified in the potato (*Solanum tuberosum* L.) SK2 gene, encoding a pistil-specific basic endochitinase (Ficker (1997) *Plant Mol. Biol.* 35:425-431). The Blec4 gene from pea (*Pisum sativum cv. Alaska*) is active in epidermal tissue of vegetative and floral shoot apices of transgenic alfalfa, making it a useful tool to target the expression of foreign genes to the epidermal layer of actively growing shoots. The activity of the Blec4 promoter in the epidermis of the shoot apex makes it particularly suitable for genetically engineering defense against insects and diseases that attack the growing shoot apex (Mandaci (1997) *Plant Mol Biol.* 34:961-965).

The invention also provides for use of tissue-specific plant promoters include a promoter from the ovule-specific *BEL1* gene described in Reiser (1995) *Cell* 83:735-742, GenBank No. U39944. Suitable seed specific promoters are derived from the following genes: *MAC1* from maize, Sheridan (1996) *Genetics* 142:1009-1020; *Cat3* from maize, GenBank No. L05934, Abler (1993) *Plant Mol. Biol.* 22:10131-1038; the gene encoding oleosin 18kD from maize, GenBank No. J05212, Lee (1994) *Plant Mol. Biol.* 26:1981-1987; vivparous-1 from *Arabidopsis,* Genbank No. U93215; the gene encoding oleosin from *Arabidopsis,* Genbank No. Z17657; Atmycl from *Arabidopsis,* Urao (1996) *Plant Mol. Biol.* 32:571-576; the 2s seed storage protein gene family from *Arabidopsis,* Conceicao (1994) *Plant* 5:493-505; the gene encoding oleosin 20kD from *Brassica napus,* GenBank No. M63985; *napA* from *Brassica napus,* GenBank No. J02798, Josefsson (1987) *JBL* 26:12196-1301; the napin gene family from *Brassica napus*, Sjodahl (1995) *Planta* 197:264-271; the gene encoding the 2S storage protein from *Brassica napus,* Dasgupta (1993) *Gene* 133:301-302; the genes encoding oleosin a, Genbank No. U09118, and, oleosin B, Genbank No. U09119, from soybean; and, the gene encoding low molecular weight sulphur rich protein from soybean, Choi (1995) *Mol Gen, Genet.* 246:266-268. The tissue specific E8 promoter from tomato is particularly useful for directing gene expression so that a desired gene product is located in fruits. Other suitable promoters include those from genes encoding embryonic storage proteins.

One of skill will recognize that a tissue-specific promoter may drive expression of operably linked sequences in tissues other than the target tissue. Thus, as

used herein a tissue-specific promoter is one that drives expression preferentially in the target tissue, but may also lead to some expression in other tissues as well.

The invention also provides for use of tissue-specific promoters derived from viruses which can include, *e.g.*, the tobamovirus subgenomic promoter (Kumagai (1995) *Proc. Natl. Acad. Sci. USA* 92:1679-1683; the rice tungro bacilliform virus (RTBV), which replicates only in phloem cells in infected rice plants, with its promoter which drives strong phloem-specific reporter gene expression; the cassava vein mosaic virus (CVMV) promoter, with highest activity in vascular elements, in leaf mesophyll cells, and in root tips (Verdaguer (1996) *Plant Mol. Biol.* 31:1129-1139).

In some embodiments, the nucleic acid construct will comprise a promoter functional in a specific plant cell, such as in a species of *Lactuca*, operably linked to an RG polynucleotide. Promoters useful in these embodiments include RG promoters. In additional embodiments, the nucleic acid construct will comprise a RG promoter operably linked to a heterologous polynucleotide. The heterologous polynucleotide is chosen to provide a plant with a desired phenotype. For example, the heterologous polynucleotide can be a structural gene which encodes a polypeptide which imparts a desired resistance phenotype. Alternatively, the heterologous polynucleotide may be a regulatory gene which might play a role in transcriptional and/or translational control to suppress, enhance, or otherwise modify the transcription and/or expression of an endogenous gene within the plant. The heterologous polynucleotide of the nucleic acid construct of the present invention can be expressed in either sense or anti-sense orientation as desired. It will be appreciated that control of gene expression in either sense or anti-sense orientation can have a direct impact on the observable plant characteristics.

Modifying and Inhibiting RG Gene Expression

The invention also provides for RG nucleic acid sequences which are complementary to the RG polypeptide-encoding sequences of the invention; *i.e.*, antisense RG nucleic acids. Antisense technology can be conveniently used to modify gene expression in plants. To accomplish this, a nucleic acid segment from the desired gene is cloned and operably linked to a promoter such that the anti-sense strand of RNA will be transcribed. The construct is then transformed into plants and the antisense strand of RNA is produced. In plant cells, it has been shown that antisense RNA inhibits gene expression by preventing the accumulation of mRNA which encodes the enzyme of interest, see, e.g.,

Sheehy (1988) *Proc. Nat. Acad. Sci. USA* 85:8805-8809; Hiatt et al., U.S. Patent No. 4,801,340.

Antisense sequences are capable of inhibiting the transport, splicing or transcription of RG-encoding genes. The inhibition can be effected through the targeting of genomic DNA or messenger RNA. The transcription or function of targeted nucleic acid can be inhibited, *e.g.*, by hybridization and/or cleavage. One particularly useful set of inhibitors provided by the present invention includes oligonucleotides which are able to either bind RG gene or message, in either case preventing or inhibiting the production or function of RG. The association can be though sequence specific hybridization. Such inhibitory nucleic acid sequences can, for example, be used to completely inhibit a plant disease resistance response. Another useful class of inhibitors includes oligonucleotides which cause inactivation or cleavage of RG message. The oligonucleotide can have enzyme activity which causes such cleavage, such as ribozymes. The oligonucleotide can be chemically modified or conjugated to an enzyme or composition capable of cleaving the complementary nucleic acid. One may screen a pool of many different such oligonucleotides for those with the desired activity.

Antisense Oligonucleotides

The invention provides for with antisense oligonucleotides capable of binding RG message which can inhibit RG activity by targeting mRNA. Strategies for designing antisense oligonucleotides are well described in the scientific and patent literature, and the skilled artisan can design such RG oligonucleotides using the novel reagents of the invention. In some situations, naturally occurring nucleic acids used as antisense oligonucleotides may need to be relatively long (18 to 40 nucleotides) and present at high concentrations. A wide variety of synthetic, non-naturally occurring nucleotide and nucleic acid analogues are known which can address this potential problem. For example, peptide nucleic acids (PNAs) containing non-ionic backbones, such as N-(2-aminoethyl) glycine units can be used. Antisense oligonucleotides having phosphorothioate linkages can also be used, as described in WO 97/03211; WO 96/39154; Mata (1997) *Toxicol Appl Pharmacol* 144:189-197; Antisense Therapeutics, ed. Agrawal (Humana Press, Totowa, N.J., 1996). Antisense oligonucleotides having synthetic DNA backbone analogues provided by the invention can also include phosphoro-dithioate, methylphosphonate,

phosphoramidate, alkyl phosphotriester, sulfamate, 3'-thioacetal, methylene(methylimino), 3'-N-carbamate, and morpholino carbamate nucleic acids, as described herein.

Combinatorial chemistry methodology can be used to create vast numbers of oligonucleotides that can be rapidly screened for specific oligonucleotides that have appropriate binding affinities and specificities toward any target, such as the sense and antisense RG sequences of the invention (for general background information, see, *e.g.*, Gold (1995) *J. of Biol. Chem.* 270:13581-13584).

Inhibitory Ribozymes

The invention provides for with ribozymes capable of binding RG message which can inhibit RG activity by targeting mRNA. Strategies for designing ribozymes and selecting the RG-specific antisense sequence for targeting are well described in the scientific and patent literature, and the skilled artisan can design such RG ribozymes using the novel reagents of the invention. Ribozymes act by binding to a target RNA through the target RNA binding portion of a ribozyme which is held in close proximity to an enzymatic portion of the RNA that cleaves the target RNA. Thus, the ribozyme recognizes and binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cleave and inactivate the target RNA. Cleavage of a target RNA in such a manner will destroy its ability to direct synthesis of an encoded protein if the cleavage occurs in the coding sequence, or, preventing transport of the message from the nucleus to the cytoplasm. After a ribozyme has bound and cleaved its RNA target, it is typically released from that RNA and so can bind and cleave new targets repeatedly.

Catalytic RNA molecules or ribozymes can also be used to inhibit expression of any plant gene. It is possible to design ribozymes that specifically pair with virtually any target RNA and cleave the phosphodiester backbone at a specific location, thereby functionally inactivating the target RNA. In carrying out this cleavage, the ribozyme is not itself altered, and is thus capable of recycling and cleaving other molecules, making it a true enzyme. The inclusion of ribozyme sequences within antisense RNAs confers RNA-cleaving activity upon them, thereby increasing the activity of the constructs. The design and use of target RNA-specific ribozymes is described, *e.g.*, in Haseloff (1988) *Nature* 334:585-591.

In some circumstances, the enzymatic nature of a ribozyme can be advantageous over other technologies, such as antisense technology (where a nucleic acid

molecule simply binds to a nucleic acid target to block its transcription, translation or association with another molecule) as the effective concentration of ribozyme necessary to effect a therapeutic treatment can be lower than that of an antisense oligonucleotide. This potential advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, a ribozyme is typically a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding, but also on the mechanism by which the molecule inhibits the expression of the RNA to which it binds. That is, the inhibition is caused by cleavage of the RNA target and so specificity is defined as the ratio of the rate of cleavage of the targeted RNA over the rate of cleavage of non-targeted RNA. This cleavage mechanism is dependent upon factors additional to those involved in base pairing. Thus, the specificity of action of a ribozyme can be greater than that of antisense oligonucleotide binding the same RNA site.

The enzymatic ribozyme RNA molecule can be formed in a hammerhead motif, but may also be formed in the motif of a hairpin, hepatitis delta virus, group I intron or RNaseP-like RNA (in association with an RNA guide sequence). Examples of such hammerhead motifs are described by Rossi (1992) *Aids Research and Human Retroviruses* 8:183; hairpin motifs by Hampel (1989) *Biochemistry* 28:4929, and Hampel (1990) *Nuc. Acids Res.* 18:299; the hepatitis delta virus motif by Perrotta (1992) *Biochemistry* 31:16; the RNaseP motif by Guerrier-Takada (1983) *Cell* 35:849; and the group I intron by Cech U.S. Pat. No. 4,987,071. The recitation of these specific motifs is not intended to be limiting; those skilled in the art will recognize that an enzymatic RNA molecule of this invention has a specific substrate binding site complementary to one or more of the target gene RNA regions, and has nucleotide sequence within or surrounding that substrate binding site which imparts an RNA cleaving activity to the molecule.

Sense Supression

Another method of suppression is sense suppression. Introduction of nucleic acid configured in the sense orientation has been shown to be an effective means by which to block the transcription of target genes. For an example of the use of this method to modulate expression of endogenous genes see, Napoli et al., *The Plant Cell* 2:279-289 (1990), and U.S. Patent No. 5,034,323.

Cloning of RG Polypeptides

Synthesis and/or cloning of RG polynucleotides and isolated nucleic acid constructs of the present invention are provided by methods well known to those of ordinary skill in the art. Generally, the nomenclature and the laboratory procedures in recombinant DNA technology described below are those well known and commonly employed in the art. Standard techniques are used for cloning, DNA and RNA isolation, amplification and purification. Generally enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like are performed according to the manufacturer's specifications. These techniques and various other techniques are generally performed according to Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, (1989).

The isolation of RG genes may be accomplished by a number of techniques. For instance, oligonucleotide probes based on the sequences disclosed here can be used to identify the desired gene in a cDNA or genomic DNA library. To construct genomic libraries, large segments of genomic DNA are generated by random fragmentation, e.g. using restriction endonucleases, and are ligated with vector DNA to form concatemers that can be packaged into the appropriate vector. To prepare a cDNA library, mRNA is isolated from the desired organ, such as roots and a cDNA library which contains the *RG* gene transcript is prepared from the mRNA. Alternatively, cDNA may be prepared from mRNA extracted from other tissues in which RG genes or homologs are expressed.

The cDNA or genomic library can then be screened using a probe based upon the sequence of a cloned RG gene such as the genes disclosed herein. Probes may be used to hybridize with genomic DNA or cDNA sequences to isolate homologous genes in the same or different plant species.

Those of skill in the art will appreciate that various degrees of stringency of hybridization can be employed in the assay; and either the hybridization or the wash medium can be stringent. As the conditions for hybridization become more stringent, there must be a greater degree of complementarity between the probe and the target for duplex formation to occur. The degree of stringency can be controlled by temperature, ionic strength, pH and the presence of a partially denaturing solvent such as formamide. For example, the stringency of hybridization is conveniently varied by changing the polarity of the reactant solution through manipulation of the concentration of formamide within the range of 0% to 50%.

Alternatively, the RG nucleic acids of the invention can be amplified from nucleic acid samples using a variety of amplification techniques, such as polymerase chain reaction (PCR) technology, to amplify the sequences of the RG and related genes directly from genomic DNA, from cDNA, from genomic libraries or cDNA libraries. PCR and other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of the desired mRNA in samples, for nucleic acid sequencing, or for other purposes.

Oligonucleotides can be used to identify and detect additional RG families and RG family species using a variety of hybridization techniques and conditions. Suitable amplification methods include, but are not limited to: polymerase chain reaction, PCR (PCR PROTOCOLS, A GUIDE TO METHODS AND APPLICATIONS, *ed.* Innis, Academic Press, N.Y. (1990) and PCR STRATEGIES (1995), *ed.* Innis, Academic Press, Inc., N.Y. (Innis)), ligase chain reaction (LCR) (Wu (1989) *Genomics* 4:560; Landegren (1988) *Science* 241:1077; Barringer (1990) *Gene* 89:117); transcription amplification (Kwoh (1989) *Proc. Natl. Acad. Sci. USA* 86:1173); and, self-sustained sequence replication (Guatelli (1990) *Proc. Natl. Acad. Sci. USA,* 87:1874); Q Beta replicase amplification and other RNA polymerase mediated techniques (*e.g.*, NASBA, Cangene, Mississauga, Ontario); see Berger (1987) *Methods Enzymol.* 152:307-316, Sambrook, and Ausubel, as well as Mullis (1987) U.S. Patent Nos. 4,683,195 and 4,683,202; Arnheim (1990) *C&EN* 36-47; Lomell *J. Clin. Chem.*, 35:1826 (1989); Van Brunt, *Biotechnology*, 8:291-294 (1990); Wu (1989) *Gene* 4:560; Sooknanan (1995) *Biotechnology* 13:563-564. Methods for cloning *in vitro* amplified nucleic acids are described in Wallace, U.S. Pat. No. 5,426,039.

The degree of complementarity (sequence identity) required for detectable binding will vary in accordance with the stringency of the hybridization medium and/or wash medium. The degree of complementarity will optimally be 100 percent; however, it should be understood that minor sequence variations in the probes and primers may be compensated for by reducing the stringency of the hybridization and/or wash medium as described earlier.

In some preferred embodiments, members of this class of pest resistance genes can be identified by their ability to be amplified by PCR primers based on the sequences disclosed here. Appropriate primers and probes for identifying RG sequences

from plant tissues are generated from comparisons of the sequences provided herein. See, e.g., Table 1. For a general overview of PCR see *PCR Protocols: A Guide to Methods and Applications*. (Innis, M, Gelfand, D., Sninsky, J. and White, T., eds.), *Academic Press*, San Diego (1990), incorporated herein by reference.

Briefly, the first step of each cycle of the PCR involves the separation of the nucleic acid duplex formed by the primer extension. Once the strands are separated, the next step in PCR involves hybridizing the separated strands with primers that flank the target sequence. The primers are then extended to form complementary copies of the target strands. For successful PCR amplification, the primers are designed so that the position at which each primer hybridizes along a duplex sequence is such that an extension product synthesized from one primer, when separated from the template (complement), serves as a template for the extension of the other primer. The cycle of denaturation, hybridization, and extension is repeated as many times as necessary to obtain the desired amount of amplified nucleic acid.

In the preferred embodiment of the PCR process, strand separation is achieved by heating the reaction to a sufficiently high temperature for an sufficient time to cause the denaturation of the duplex but not to cause an irreversible denaturation of the polymerase (see U.S. Patent No. 4,965,188). Template-dependent extension of primers in PCR is catalyzed by a polymerizing agent in the presence of adequate amounts of four deoxyribonucleotide triphosphates (typically dATP, dGTP, dCTP, and dTTP) in a reaction medium comprised of the appropriate salts, metal cations, and pH buffering system. Suitable polymerizing agents are enzymes known to catalyze template-dependent DNA synthesis.

Polynucleotides may also be synthesized by well-known techniques as described in the technical literature. See, e.g., Carruthers *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 47:411-418 (1982), and Adams *et al.*, *J. Am. Chem. Soc.* 105:661 (1983). Double stranded DNA fragments may then be obtained either by synthesizing the complementary strand and annealing the strands together under appropriate conditions, or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

RG Proteins

The present invention further provides isolated RG proteins encoded by the RG polynucleotides disclosed herein. One of skill will recognize that the nucleic acid encoding a functional RG protein need not have a sequence identical to the exemplified genes disclosed here. For example, because of codon degeneracy a large number of nucleic acid sequences can encode the same polypeptide. In addition, the polypeptides encoded by the RG genes, like other proteins, have different domains which perform different functions. Thus, the RG gene sequences need not be full length, so long as the desired functional domain of the protein is expressed.

The resistance proteins are at least 25 amino acid residues in length. Typically, the RG proteins are at least 50 amino acid residues, generally at least 100, preferably at least 150, more preferably at least 200 amino acids in length. In particularly preferred embodiments, the RG proteins are of sufficient length to provide resistance to pests when expressed in the desired plants. Generally then, the RG proteins will be the length encoded by an RG gene of the present invention. However, those of ordinary skill will appreciate that minor deletions, substitutions, or additions to an RG protein will typically yield a protein with pest resistance characteristics similar or identical to that of the full length sequence. Thus, full-length RG proteins modified by 1, 2, 3, 4, or 5 deletions, substitutions, or additions, generally provide an effective degree of pest resistance relative to the full-length protein.

The RG proteins which provide pest resistance will typically comprise at least one of an LRR or an NBS. Preferably, both are present. LRR and/or NBS regions present in the RG proteins of the present invention can be provided by RG genes of the present invention. In some embodiments, the LRR and/or NBS regions are obtained from other pest resistance genes. See, e.g., Yu *et al., Proc. Natl. Acad. Sci. USA*, 93: 11751-11756 (1996); Bent *et al.*, *Science*, 265: 1856-1860 (1994).

Modified protein chains can also be readily designed utilizing various recombinant DNA techniques well known to those skilled in the art. For example, the chains can vary from the naturally occurring sequence at the primary structure level by amino acid substitutions, additions, deletions, and the like. Modification can also include swapping domains from the proteins of the invention with related domains from other pest resistance genes.

Pests that can be targeted by RG genes and proteins of the present invention include such bacterial pests as *Erwinia carotovora* and *Pseudomonas marginalis*. Fungal pests which can be targeted by the present invention include *Bremia lactucae*, *Marssonina panattoniana*, *Rhizoctonia solani*, *Olpidium brassicae*, root aphid, *Sclerotinia sclerotiorum* and *S. minor*, and *Botrytis cinerea* which causes gray mold. RG genes also provide resistance to viral diseases such as lettuce and turnip mosaic viruses.

Fusion Proteins

RG polypeptides can also be expressed as recombinant proteins with one or more additional polypeptide domains linked thereto to facilitate protein detection, purification, or other applications. Such detection and purification facilitating domains include, but are not limited to, metal chelating peptides such as polyhistidine tracts and histidine-tryptophan modules that allow purification on immobilized metals, protein a domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle WA). The inclusion of a cleavable linker sequences such as Factor Xa or enterokinase (Invitrogen, San Diego CA) between the purification domain and plant disease resistant polypeptide may be useful to facilitate purification. One such expression vector provides for expression of a fusion protein comprising the sequence encoding a plant disease resistant polypeptide of the invention and nucleic acid sequence encoding six histidine residues followed by thioredoxin and an enterokinase cleavage site (*e.g.*, see Williams (1995) *Biochemistry* 34:1787-1797). The histidine residues facilitate detection and purification while the enterokinase cleavage site provides a means for purifying the desired protein(s) from the remainder of the fusion protein. Technology pertaining to vectors encoding fusion proteins and application of fusion proteins are well described, see *e.g.*, Kroll (1993) *DNA Cell. Biol.*, 12:441-53.

Antibodies Reactive to RG Polypeptides and Immunological Assays

The present invention also provides antibodies which specifically react with RG proteins of the present invention under immunologically reactive conditions. An antibody immunologically reactive with a particular antigen can be generated *in vivo* or by recombinant methods such as selection of libraries of recombinant antibodies in phage or similar vectors. "Immunologically reactive conditions" includes reference to conditions which allow an antibody, generated to a particular epitope of an antigen, to bind to that

epitope to a detectably greater degree than the antibody binds to substantially all other epitopes, generally at least two times above background binding, preferably at least five times above background. Immunologically reactive conditions are dependent upon the format of the antibody binding reaction and typically are those utilized in immunoassay protocols.

"Antibody" includes reference to an immunoglobulin molecule obtained by *in vitro* or *in vivo* generation of the humoral response, and includes both polyclonal and monoclonal antibodies. The term also includes genetically engineered forms such as chimeric antibodies (e.g., humanized murine antibodies), heteroconjugate antibodies (e.g., bispecific antibodies), and recombinant single chain Fv fragments (scFv). The term "antibody" also includes antigen binding forms of antibodies (e.g., Fab', $F(ab')_2$, Fab, Fv, rIgG, and, inverted IgG). See, Pierce Catalog and Handbook, 1994-1995 (Pierce Chemical Co., Rockford, IL). An antibody immunologically reactive with a particular antigen can be generated *in vivo* or by recombinant methods such as selection of libraries of recombinant antibodies in phage or similar vectors. *See, e.g.,* Huse *et al.* (1989) *Science* 246:1275-1281; and Ward, *et al.* (1989) *Nature* 341:544-546; and Vaughan *et al.* (1996) *Nature Biotechnology*, 14:309-314.

Many methods of making antibodies are known to persons of skill. A number of immunogens are used to produce antibodies specifically reactive to an isolated RG protein of the present invention under immunologically reactive conditions. An isolated recombinant, synthetic, or native RG protein of the present invention is the preferred immunogens (antigen) for the production of monoclonal or polyclonal antibodies.

The RG protein is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies can be generated for subsequent use in immunoassays to measure the presence and quantity of the RG protein. Methods of producing monoclonal or polyclonal antibodies are known to those of skill in the art. *See, e.g.,* Coligan (1991) *Current Protocols in Immunology* Wiley/Greene, NY; and Harlow and Lane (1989) *Antibodies: A Laboratory Manual* Cold Spring Harbor Press, NY); Goding (1986) *Monoclonal Antibodies: Principles and Practice* (2d ed.) Academic Press, New York, NY.

Frequently, the RG proteins and antibodies will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide

variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionucleotides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, chemiluminescent moieties, magnetic particles, and the like. Patents teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

The antibodies of the present invention can be used to screen plants for the expression of RG proteins of the present invention. The antibodies of this invention are also used for affinity chromatography in isolating RG protein.

The present invention further provides RG polypeptides that specifically bind. under immunologically reactive conditions, to an antibody generated against a defined immunogen, such as an immunogen consisting of the RG polypeptides of the present invention. Immunogens will generally be at least 10 contiguous amino acids from an RG polypeptide of the present invention. Optionally, immunogens can be from regions exclusive of the NBS and/or LRR regions of the RG polypeptides. Nucleic acids which encode such cross-reactive RG polypeptides are also provided by the present invention. The RG polypeptides can be isolated from any number plants as discussed earlier. Preferred are species from the family *Compositae* and in particular the genus *Lactuca* such as *L. sativa* and such subspecies as *crispa, longifolia,* and *asparagina*.

"Specifically binds" includes reference to the preferential association of a ligand, in whole or part, with a particular target molecule (i.e., "binding partner" or "binding moiety") relative to compositions lacking that target molecule. It is, of course, recognized that a certain degree of non-specific interaction may occur between a ligand and a non-target molecule. Nevertheless, specific binding, may be distinguished as mediated through specific recognition of the target molecule. Typically specific binding results in a much stronger association between the ligand and the target molecule than between the ligand and non-target molecule. Specific binding by an antibody to a protein under such conditions requires an antibody that is selected for its specificity for a particular protein. The affinity constant of the antibody binding site for its cognate monovalent antigen is at least 10^7, usually at least 10^8, preferably at least 10^9, more preferably at least 10^{10}, and most preferably at least 10^{11} liters/mole. A variety of immunoassay formats are appropriate for selecting antibodies specifically reactive with a particular protein. For

example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically reactive with a protein. See Harlow and Lane (1988) Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific reactivity. The antibody may be polyclonal but preferably is monoclonal. Generally, antibodies cross-reactive to such proteins as RPS2, RPM1 (bacterial resistances in Arabidopsis, L6 (fungal resistance in flax, PRF (resistance to *Pseudomonas syringae* in tomator), and *N,* (virus resistance in tobacco), are removed by immunoabsorbtion.

Immunoassays in the competitive binding format are typically used for cross-reactivity determinations. For example, an immunogenic RG polypeptide is immobilized to a solid support. Polypeptides added to the assay compete with the binding of the antisera to the immobilized antigen. The ability of the above polypeptides to compete with the binding of the antisera to the immobilized RG polypeptide is compared to the immunogenic RG polypeptide. The percent cross-reactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% cross-reactivity with such proteins as RPS2, RPM1, L6, PRF, and *N,* are selected and pooled. The cross-reacting antibodies are then removed from the pooled antisera by immunoabsorbtion with these non-RG resistance proteins.

The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay to compare a second "target" polypeptide to the immunogenic polypeptide. In order to make this comparison, the two polypeptides are each assayed at a wide range of concentrations and the amount of each polypeptide required to inhibit 50% of the binding of the antisera to the immobilized protein is determined using standard techniques. If the amount of the target polypeptide required is less than twice the amount of the immunogenic polypeptide that is required, then the target polypeptide is said to specifically bind to an antibody generated to the immunogenic protein. As a final determination of specificity, the pooled antisera is fully immunosorbed with the immunogenic polypeptide until no binding to the polypeptide used in the immunosorbtion is detectable. The fully immunosorbed antisera is then tested for reactivity with the test polypeptide. If no reactivity is observed, then the test polypeptide is specifically bound by the antisera elicited by the immunogenic protein.

Production of transgenic plants of the invention

Isolated nucleic acid constructs prepared as described herein can be introduced into plants according techniques known in the art. In some embodiments, the introduced nucleic acid is used to provide RG gene expression and therefore pest resistance in desired plants. In some embodiments, RG promoters are used to drive expression of desired heterologous genes in plants. Finally, in some embodiments, the constructs can be used to suppress expression of a target endogenous gene, including RG genes.

To use isolated RG sequences in the above techniques, recombinant DNA vectors suitable for transformation of plant cells are prepared. Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, for example, Weising *et al. Ann. Rev. Genet.* 22:421-477 (1988).

A DNA sequence coding for the desired RG polypeptide, for example a cDNA or a genomic sequence encoding a full length protein, will be used to construct a recombinant expression cassette which can be introduced into the desired plant. An expression cassette will typically comprise the RG polynucleotide operably linked to transcriptional and translational initiation regulatory sequences which will direct the transcription of the sequence from the RG gene in the intended tissues of the transformed plant.

Such DNA constructs may be introduced into the genome of the desired plant host by a variety of conventional techniques. For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation, PEG poration, particle bombardment and microinjection of plant cell protoplasts or embryogenic callus, or the DNA constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment. Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria.

Transformation techniques are known in the art and well described in the scientific and patent literature. The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski *et al. Embo J.* 3:2717-2722 (1984).

Electroporation techniques are described in Fromm *et al. Proc. Natl. Acad. Sci. USA* 82:5824 (1985). Ballistic transformation techniques are described in Klein *et al. Nature* 327:70-73 (1987).

Agrobacterium tumefaciens-meditated transformation techniques are well described in the scientific literature. See, for example Horsch *et al. Science* 233:496-498 (1984), and Fraley *et al. Proc. Natl. Acad. Sci. USA* 80:4803 (1983). Although *Agrobacterium* is useful primarily in dicots, certain monocots can be transformed by *Agrobacterium*. For instance, *Agrobacterium* transformation of rice is described by Hiei *et al*, *Plant J.* 6:271-282 (1994). A particularly preferred means of transforming lettuce is described in Michelmore *et al.*, *Plant Cell Reports*, 6:439-442 (1987).

Transformed plant cells which are derived by any of the above transformation techniques can be cultured to regenerate a whole plant which possesses the transformed genotype and thus the desired RG-controlled phenotype. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker which has been introduced together with the RG nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans *et al.*, *Protoplasts Isolation and Culture*, *Handbook of Plant Cell Culture*, pp. 124-176, Macmillilan Publishing Company, New York, 1983; and Binding, *Regeneration of Plants*, *Plant Protoplasts*, pp. 21-73, CRC Press, Boca Raton, 1985. Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee *et al. Ann. Rev. of Plant Phys.* 38:467-486 (1987).

The methods of the present invention are particularly useful for incorporating the RG polynucleotides into transformed plants in ways and under circumstances which are not found naturally. In particular, the RG polypeptides may be expressed at times or in quantities which are not characteristic of natural plants.

One of skill will recognize that after the expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used. depending upon the species to be crossed.

<u>Detection of RG Resistance Genes</u>

The present invention further provides methods for detecting RG resistance genes in a nucleic acid sample suspected of comprising an RG resistance gene. The means by which the RG resistance gene is detected is not a critical aspect of the invention. For example, RG resistance genes can be detected by the presence of amplicons using RG resistance gene specific primers. Additionally, RG resistance genes can be detected by assaying for specific hybridization of an RG polynucleotide to an RG resistance gene. In some embodiments, the RG resistance gene can be amplified prior to the step of contacting the nucleic acid sample with the RG polynucleotide.

In a typical detection method, the nucleic acid sample is contacted with an RG polynucleotide to form a hybridization complex. The hybridization complex may be detected directly (e.g., in Southern or northern blots), or indirectly (e.g., by subsequent primer extension during PCR amplification). The RG polynucleotide hybridizes under stringent conditions to an RG polynucleotide of the invention. Formation of the hybridization complex is directly or indirectly used to indicate the presence of the RG resistance gene in the nucleic acid sample.

Detection of the hybridization complex can be achieved using any number of well known methods. For example, the nucleic acid sample, or a portion thereof, may be assayed by hybridization formats including but not limited to, solution phase, solid phase, mixed phase, or *in situ* hybridization assays. Briefly, in solution (or liquid) phase hybridizations, both the target nucleic acid and the probe or primer are free to interact in the reaction mixture. In solid phase hybridization assays, probes or primers are typically linked to a solid support where they are available for hybridization with target nucleic in solution. In mixed phase, nucleic acid intermediates in solution hybridize to target nucleic acids in solution as well as to a nucleic acid linked to a solid support. In *in situ* hybridization, the target nucleic acid is liberated from its cellular surroundings in such as to be available for hybridization within the cell while preserving the cellular morphology for subsequent interpretation and analysis. The following articles provide an overview of the various hybridization assay formats: Singer *et al., Biotechniques 4(3)*:230-250 (1986); Haase *et al., Methods in Virology*, Vol. VII, pp. 189-226 (1984); Wilkinson, "The theory and practice of in situ hybridization" In: *In situ Hybridization*, Ed. D.G. Wilkinson. IRL Press, Oxford University Press, Oxford; and *Nucleic Acid Hybridization: A Practical Approach*, Ed. Hames, B.D. and Higgins, S.J., IRL Press (1987).

The effect of the modification of RG gene expression can be measured by detection of increases or decreases in mRNA levels using, for instance, Northern blots. In addition, the phenotypic effects of gene expression can be detected by measuring nematode, fungal, bacterial, viral, or other pest resistance in plants. Suitable assays for determining pest resistance are well known. Michelmore and Crute, *Trans. Br. mycol. Soc,* 79(3): 542-546 (1982).

The means by which hybridization complexes are detected is not a critical aspect of the present invention and can be accomplished by any number of methods currently known or later developed. RG polynucleotides can be labeled by any one of several methods typically used to detect the presence of hybridized nucleic acids. One common method of detection is the use of autoradiography using probes labeled with ^{3}H, ^{125}I, ^{35}S, ^{14}C, or ^{32}P, or the like. The choice of radioactive isotope depends on research preferences due to ease of synthesis, stability, and half lives of the selected isotopes. Other labels include ligands which bind to antibodies labeled with fluorophores, chemiluminescent agents, and enzymes. Alternatively, probes can be conjugated directly with labels such as fluorophores, chemiluminescent agents or enzymes. The choice of label depends on sensitivity required, ease of conjugation with the probe, stability requirements, and available instrumentation. Labeling the RG polynucleotide is readily achieved such as by the use of labeled PCR primers.

The choice of label dictates the manner in which the label is bound to the probe. Radioactive probes are typically made using commercially available nucleotides containing the desired radioactive isotope. The radioactive nucleotides can be incorporated into probes, for example, by using DNA synthesizers, by nick translation with DNA polymerase I, by tailing radioactive DNA bases to the 3' end of probes with terminal deoxynucleotidyl transferase, by treating single-stranded M13 plasmids having specific inserts with the Klenow fragment of DNA polymerase in the presence of radioactive deoxynucleotides, dNTP, by transcribing from RNA templates using reverse transcriptase in the presence of radioactive deoxynucleotides, dNTP, or by transcribing RNA from vectors containing specific RNA viral promoters (e.g., SP6 promoter) using the corresponding RNA polymerase (e.g., SP6 RNA polymerase) in the presence of radioactive ribonucleotides rNTP.

The probes can be labeled using radioactive nucleotides in which the isotope resides as a part of the nucleotide molecule, or in which the radioactive component is attached to the nucleotide via a terminal hydroxyl group that has been esterified to a radioactive component such as inorganic acids, *e.g.*, 32P phosphate or 14C organic acids, or esterified to provide a linking group to the label. Base analogs having nucleophilic linking groups, such as primary amino groups, can also be linked to a label.

Non-radioactive probes are often labeled by indirect means. For example, a ligand molecule is covalently bound to the probe. The ligand then binds to an anti-ligand molecule which is either inherently detectable or covalently bound to a detectable signal system, such as an enzyme, a fluorophore, or a chemiluminescent compound. Enzymes of interest as labels will primarily be hydrolases, such as phosphatases, esterases and glycosidases, or oxidoreductases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescers include luciferin, and 2,3-dihydrophthalazinediones, *e.g.*, luminol. Ligands and anti-ligands may be varied widely. Where a ligand has a natural anti-ligand, namely ligands such as biotin, thyroxine, and cortisol, it can be used in conjunction with its labeled, naturally occurring anti-ligands. Alternatively, any haptenic or antigenic compound can be used in combination with an antibody.

Probes can also be labeled by direct conjugation with a label. For example, cloned DNA probes have been coupled directly to horseradish peroxidase or alkaline phosphatase, (Renz. M., and Kurz, K. (1984) A Colorimetric Method for DNA Hybridization. *Nucl. Acids Res.* 12: 3435-3444) and synthetic oligonucleotides have been coupled directly with alkaline phosphatase (Jablonski, E., *et al.* (1986) Preparation of Oligodeoxynucleotide-Alkaline Phosphatase Conjugates and Their Use as Hybridization Probes. *Nuc. Acids. Res.* 14: 6115-6128; and Li P., *et al.* (1987) Enzyme-linked Synthetic Oligonucleotide probes: Non-Radioactive Detection of Enterotoxigenic *Escherichia Coli* in Faeca Specimens. *Nucl. Acids Res.* 15:5275-5287).

Definitions

Units, prefixes, and symbols can be denoted in their SI accepted form. Numeric ranges are inclusive of the numbers defining the range. Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation, respectively. The

headings provided herein are not limitations of the various aspects or embodiments of the invention which can be had by reference to the specification as a whole. Accordingly, the terms defined immediately below are more fully defined by reference to the specification as a whole.

As used herein, the term "plant" includes reference to whole plants, plant organs (e.g., leaves, stems, roots, etc.), seeds and plant cells and progeny of same. The class of plants which can be used in the methods of the invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants.

As used herein, "pest" includes, but is not limited to, viruses, fungi, nematodes, insects, and bacteria.

As used herein, "heterologous" is a nucleic acid that originates from a foreign species, or, if from the same species, is substantially modified from its original form. For example, a promoter operably linked to a heterologous structural gene is from a species different from that from which the structural gene was derived, or, if from the same species, one or both are substantially modified from their original form.

As used herein, "RG gene," alternatively referred to as "RLG gene," is a gene encoding resistance to plant pests, such as viruses, fungi, nematodes, insects, and bacteria, and which hybridizes under stringent conditions and/or has at least 60% sequence identity at the deduced amino acid level to the exemplified sequences provided herein. RG genes encode "RG polypeptides," alternatively referred to as "RLG polypeptides," which can comprise LRR motifs and/or NBS motifs. The RG polypeptides encoded by RG genes have at least 55% or 60% sequence identity, typically at least 65% sequence identity, preferably at least 70% sequence identity, often at least 75% sequence identity, more preferably at least 80% sequence identity, and most preferably at least 90% sequence identity at the deduced amino acid level relative to the exemplary RG sequences provided herein. The term "RG family" or "RG family genus" or "genus" includes reference to a group of RG polypeptide sequence species that have at least 60% amino acid sequence identity, and, the nucleic acids encoding these polypeptides. The individual species of a genus, i.e., the members of a family, typically are genetically mapped to the same locus.

As used herein, "RG polynucleotide" includes reference to a contiguous sequence from an RG gene of at least 18, 20, 25, 30, 40, or 50 nucleotides in length, up to

at least about 100 or at least about 200 nucleotides in length. In some embodiments, the polynucleotide is preferably at least 100 nucleotides in length, more preferably at least 200 nucleotides in length, most preferably at least 500 nucleotides in length. Thus, RG polynucleotide may be a RG gene or a subsequence thereof.

As used herein, "isolated," when referring to a molecule or composition, such as, for example, an RG polypeptide or nucleic acid, means that the molecule or composition is separated from at least one other compound, such as a protein, other nucleic acids (*e.g.*, RNAs), or other contaminants with which it is associated *in vivo* or in its naturally occurring state. Thus, an RG polypeptide or nucleic acid is considered isolated when it has been isolated from any other component with which it is naturally associated, *e.g.*. cell membrane, as in a cell extract. An isolated composition can, however, also be substantially pure. An isolated composition can be in a homogeneous state and can be in a dry or an aqueous solution. Purity and homogeneity can be determined, for example, using analytical chemistry techniques such as polyacrylamide gel electrophoresis (SDS-PAGE) or high performance liquid chromatography (HPLC).

The term "nucleic acid" or “nucleic acid molecule” or “nucleic acid sequence” refers to a deoxyribonucleotide or ribonucleotide oligonucleotide in either single- or double-stranded form. The term encompasses nucleic acids, *i.e.*, oligonucleotides, containing known analogues of natural nucleotides which have similar or improved binding properties, for the purposes desired, as the reference nucleic acid. The term also includes nucleic acids which are metabolized in a manner similar to naturally occurring nucleotides or at rates that are improved thereover for the purposes desired. The term also encompasses nucleic-acid-like structures with synthetic backbones. DNA backbone analogues provided by the invention include phosphodiester, phosphorothioate, phosphorodithioate, methylphosphonate, phosphoramidate, alkyl phosphotriester, sulfamate, 3'-thioacetal, methylene(methylimino), 3'-N-carbamate, morpholino carbamate, and peptide nucleic acids (PNAs); see Oligonucleotides and Analogues, a Practical Approach, edited by F. Eckstein, IRL Press at Oxford University Press (1991); Antisense Strategies, Annals of the New York Academy of Sciences, Volume 600, Eds. Baserga and Denhardt (NYAS 1992); Milligan (1993) *J. Med. Chem.* 36:1923-1937; Antisense Research and Applications (1993, CRC Press). PNAs contain non-ionic backbones, such as N-(2-aminoethyl) glycine units. Phosphorothioate linkages are described in WO

97/03211; WO 96/39154; Mata (1997) *Toxicol Appl Pharmacol* 144:189-197. Other synthetic backbones encompasses by the term include methyl-phosphonate linkages or alternating methylphosphonate and phosphodiester linkages (Strauss-Soukup (1997) *Biochemistry* 36:8692-8698), and benzylphosphonate linkages (Samstag (1996) *Antisense Nucleic Acid Drug Dev* 6:153-156). The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide primer, probe and amplification product. Unless otherwise indicated, a particular nucleic acid sequence includes the complementary sequence thereof.

The term "exogenous nucleic acid" refers to a nucleic acid that has been isolated, synthesized, cloned, ligated, excised in conjunction with another nucleic acid, in a manner that is not found in nature, and/or introduced into and/or expressed in a cell or cellular environment other than or at levels or forms different than the cell or cellular environment in which said nucleic acid or protein is be found in nature. The term encompasses both nucleic acids originally obtained from a different organism or cell type than the cell type in which it is expressed, and also nucleic acids that are obtained from the same cell line as the cell line in which it is expressed. invention.

The term "recombinant," when used with reference to a cell, or to the nucleic acid, protein or vector refers to a material, or a material corresponding to the natural or native form of the material, that has been modified by the introduction of a new moiety or alteration of an existing moiety, or is identical thereto but produced or derived from synthetic materials. For example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise expressed at a different level, typically, under-expressed or not expressed at all. The term "recombinant means" encompasses all means of expressing, *i.e.*, transcription or translation of, an isolated and/or cloned nucleic acid *in vitro* or *in vivo*. For example, the term "recombinant means" encompasses techniques where a recombinant nucleic acid, such as a cDNA encoding a protein, is inserted into an expression vector, the vector is introduced into a cell and the cell expresses the protein. "Recombinant means" also encompass the ligation of nucleic acids having coding or promoter sequences from different sources into one vector for expression of a fusion protein, constitutive expression of a protein, or inducible expression of a protein, such as the plant disease resistant, or RG. polypeptides of the invention.

The term "specifically hybridizes" refers to a nucleic acid that hybridizes, duplexes or binds to a particular target DNA or RNA sequence. The target sequences can be present in a preparation of total cellular DNA or RNA. Proper annealing conditions depend, for example, upon a nucleic acid's, such as a probe's length, base composition, and the number of mismatches and their position on the probe, and can be readily determined empirically providing the appropriate reagents are available. For discussions of nucleic acid probe design and annealing conditions, see, *e.g.*, Sambrook and Ausubel.

The terms "stringent hybridization," "stringent conditions," or "specific hybridization conditions" refers to conditions under which an oligonucleotide (when used, for example, as a probe or primer) will hybridize to its target subsequence, such as an RG nucleic acid in an expression vector of the invention but not to a non-RG sequence. Stringent conditions are sequence-dependent. Thus, in one set of stringent conditions an oligonucleotide probe will hybridize to only one specie of the genus of RG nucleic acids of the invention. In another set of stringent conditions (less stringent) an oligonucleotide probe will hybridize to all species of the invention's genus but not to non-RG nucleic acids. Longer sequences hybridize specifically at higher temperatures. Stringent conditions are selected to be about $5^{O}C$ lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium (if the target sequences are present in excess, at T_m, 50% of the probes are occupied at equilibrium). Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, *i.e.*, about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about $30^{O}C$ for short probes (*e.g.*, 10 to 50 nucleotides) and at least about $60^{O}C$ for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Often, high stringency wash conditions preceded by low stringency wash conditions to remove background probe signal. An example of medium stringency wash conditions for a duplex of, *e.g.*, more than 100 nucleotides, is 1x SSC at $45^{O}C$ for 15 minutes (see Sambrook for a description of SSC buffer). An example low stringency wash for a duplex of, *e.g.*, more than 100 nucleotides, is 4-6x SSC at $40^{O}C$ for 15 minutes. a signal to noise ratio of 2x (or higher) than that observed for an unrelated

probe in the particular hybridization assay indicates detection of a "specific hybridization." Nucleic acids which do not hybridize to each other under stringent conditions can still be substantially identical if the polypeptides which they encode are substantially identical. This can occurs, *e.g.*, when a nucleic acid is created that encodes for conservative substitutions. Stringent hybridization and stringent hybridization wash conditions are different under different environmental parameters, such as for Southern and Northern hybridizations. An extensive guide to the hybridization of nucleic acids is found in, *e.g.*, Sambrook, Tijssen (1993) *supra*.

As used herein "operably linked" includes reference to a functional linkage between a promoter and a second sequence, wherein the promoter sequence initiates and mediates transcription of the DNA sequence corresponding to the second sequence. Generally, operably linked means that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame.

In the expression of transgenes one of skill will recognize that the inserted polynucleotide sequence need not be identical and may be "substantially identical" to a sequence of the gene from which it was derived. As explained herein, these variants are specifically covered by this term.

In the case where the inserted polynucleotide sequence is transcribed and translated to produce a functional RG polypeptide, one of skill will recognize that because of codon degeneracy, a number of polynucleotide sequences will encode the same polypeptide. These variants are specifically covered by the term "RG polynucleotide sequence". In addition, the term specifically includes those full length sequences substantially identical (determined as described herein) with an RG gene sequence which encode proteins that retain the function of the RG protein. Thus, in the case of RG genes disclosed here, the term includes variant polynucleotide sequences which have substantial identity with the sequences disclosed here and which encode proteins capable of conferring resistance to nematodes, bacteria, viruses, fungi, insects or other pests on a transgenic plant comprising the sequence.

Two polynucleotides or polypeptides are said to be "identical" if the sequence of nucleotides or amino acid residues, respectively, in the two sequences is the same when aligned for maximum correspondence, as described below. The term

"complementary to" is used herein to mean that the complementary sequence is identical to all or a specified contiguous portion of a reference polynucleotide sequence.

The terms "sequence identity," "sequence similarity" and "homology" refer to when two sequences, such as the nucleic acid and amino acid sequences or the polypeptides of the invention, when optimally aligned, as with, for example, the programs PILEUP, BLAST, GAP, FASTA or BESTFIT (see discussion, *supra*). "Percentage amino acid/nucleic acid sequence identity" refers to a comparison of the sequences of two polypeptides/nucleic acids which, when optimally aligned, have approximately the designated percentage of the same amino acids/nucleic acids, respectively. For example, "60% sequence identity" and "60% homology" refer to a comparison of the sequences of two RG nucleic acids or polypeptides which, when optimally aligned, have 60% identity. For example, in one embodiment, nucleic acids encoding RG polypeptides of the invention comprise a sequence with at least 50% nucleic acid sequence identity to SEQ ID NO:1. In other embodiments, the RG polypeptides of the invention are encoded by nucleic acids comprising a sequence with at least 50% sequence identity to SEQ ID NO:1, or, are encoded by nucleic acids comprising SEQ ID NO:1, or, have at least 60% amino acid sequence identity to the polypeptide of SEQ ID NO:2.

"Percentage of sequence identity" is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

The term "substantial identity" of polynucleotide sequences means that a polynucleotide comprises a sequence that has at least 55% or 60% sequence identity, generally at least 65%, preferably at least 70%, often at least 75%, more preferably at least 80% and most preferably at least 90%, compared to a reference sequence using the programs described above (preferably BESTFIT) using standard parameters. One of skill will recognize that these values can be appropriately adjusted to determine corresponding

identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like. Substantial identity of amino acid sequences for these purposes normally means sequence identity of at least 55% or 60%, preferably at least 70%, more preferably at least 80%, and most preferably at least 95%. Polypeptides having "sequence similarity" share sequences as noted above except that residue positions which are not identical may differ by conservative amino acid changes. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under appropriate conditions. Appropriate conditions can be high or low stringency and will be different in different circumstances. Generally, stringent conditions are selected to be about 5°C to about 20°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically, stringent wash conditions are those in which the salt concentration is about 0.02 molar at pH 7 and the temperature is at least about 50°C. However, nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This may occur, *e.g.*, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code.

Nucleic acids of the invention can be identified from a cDNA or genomic library prepared according to standard procedures and the nucleic acids disclosed here used as a probe. Thus, for example, stringent hybridization conditions will typically include at least one low stringency wash using 0.3 molar salt (e.g., 2X SSC) at 65°C. The washes

are preferably followed by one or more subsequent washes using 0.03 molar salt (e.g., 0.2X SSC) at 50°C, usually 60°C, or mosre usually 65°C. Nucleic acid probes used to identify the nucleic acids are preferably at least 100 nucleotides in length.

As used herein, "nucleotide binding site" or "nucleotide binding domain" ("NBS") includes reference to highly conserved nucleotide-, *i.e.*, ATP/GTP-, binding domains, typically included in the "kinase domain" of kinase polypeptides, such as a kinase-1a, kinase 2, or a kinase 3a motif, as described herein. For example, the tobacco N and Arabidopsis RPS2 genes, among several recently cloned disease-resistance genes, share highly conserved NBS sequence. Kinase NBS subdomains further consist of three subdomain motifs: the P-loop, kinase-2, and kinase-3a subdomains (Yu (1996) *Proc. Acad. Sci. USA* 93:11751-11756). As discussed in detail herein, examples include the *Arabidopsis* RPP5 gene (Parker (1997) *supra*), the *A. thaliana* RPS2 gene (Mindrinos (1997) *supra*), and the flax L6 rust resistance gene (Lawrence (1995) *supra*) which all encode proteins containing an NBS; and Mindrinos (1994) *Cell* 78:1089-1099; and Shen (1993) *FEBS* 335:380-385. Using the teachings disclosed and incorporated herein and standard nucleic acid hybridization and/or amplification techniques, one of skill can identify members having NBS domains, including any of the genus of NBS-containing plant disease resistant polypeptides of the invention.

As used herein, "leucine rich region" ("LRR") includes reference to a region that has a leucine content of at least 20% leucine or isoleucine, or 30% of the aliphatic residues: leucine, isoleucine, methionine, valine, and phenylalanine, and arranged with approximate repeated periodicity. The length of the repeat may vary in length but is generally about 20 to 30 amino acids. An LRR-containing polypeptide typicially will have the canonical 24 amino acid leucine-rich repeat (LRR) sequence , which is present in different proteins that mediates molecular recognition and/or interaction processes; as described in Bent (1994) *Science* 265:1856-1860; Parker (1997) *Plant Cell.* 9:879-894; Hong (1997) *Plant Physiol.* 113:1203-1212; Schmitz (1997) *Nucleic Acids Res.* 25:756-763; Hipskind (1996) *Mol. Plant Microbe Interact.* 9:819-825; Tornero (1996) *Plant J.* 10:315-330; Dixon (1996) *Cell* 84:451-459; Jones (1994) *Science* 266:789-793; Lawrence (1995) *Plant Cell* 7:1195-1206; Song (1995) *Science* 270:1804-1806; as discussed in further detail *supra*. Using the teachings disclosed and incorporated herein and standard nucleic acid hybridization and/or amplification techniques, one of skill can

identify polypeptides having LRR domains, including any member of the genus of LRR-containing RG polypeptides of the invention.

The term "promoter" refers to a region or sequence determinants located upstream or downstream from the start of transcription and which are involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A "plant promoter" is a promoter capable of initiating and/or regulating transcription in plant cells; see also discussion on plant promoters, *supra*.

The term "constitutive promoter" refers to a promoter that initiates and helps control transcription in all tissues. Promoters that drive expression continuously under physiological conditions are referred to herein as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation; see also detailed discussion, *supra*.

The term "inducible promoter" refers to a promoter which directs transcription under the influence of changing environmental conditions or developmental conditions. Examples of environmental conditions that may effect transcription by inducible promoters include anaerobic conditions, elevated temperature, drought, or the presence of light. Such promoters are referred to herein as "inducible" promoters; see also detailed discussion, *supra*.

The term "abscission-induced promoter" or "abcission promoter" refers to a class of promoters which are activated upon plant ripening, such as fruit ripening, and are especially useful incorporated in the expression systems (*e.g.*, expression cassettes, vectors) of the invention. When the plant disease resistant polypeptide-encoding nucleic acid is under the control of an abcission promoter, rapid cell death, induced by expression of the invention's polypeptide, accelerates and/or accentuates abcission of the plant part, increasing the efficiency of the harvesting of fruits or other plant parts, such as cotton, and the like; see also detailed discussion, *supra*.

The term "tissue-specific promoter" refers to a class of transcriptional control elements that are only active in particular cells or tissues. Examples of plant promoters under developmental control include promoters that initiate transcription only (or primarily only) in certain tissues, such as roots, leaves, fruit, ovules, seeds, pollen, pistols, or flowers; see also detailed discussion, *supra*.

As used herein "recombinant" includes reference to a cell, or nucleic acid, or vector, that has been modified by the introduction of a heterologous nucleic acid or the alteration of a native nucleic acid to a form not native to that cell, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

As used herein, a "recombinant expression cassette" or "expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements which permit transcription of a particular nucleic acid in a target cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the recombinant expression cassette portion of the expression vector includes a nucleic acid to be transcribed, and a promoter.

As used herein, "transgenic plant" includes reference to a plant modified by introduction of a heterologous polynucleotide. Generally, the heterologous polynucleotide is an RG structural or regulatory gene or subsequences thereof.

As used herein, "hybridization complex" includes reference to a duplex nucleic acid sequence formed by selective hybridization of two single-stranded nucleic acids with each other.

As used herein, "amplified" includes reference to an increase in the molarity of a specified sequence. Amplification methods include the polymerase chain reaction (PCR), the ligase chain reaction (LCR), the transcription-based amplification system (TAS), the self-sustained sequence replication system (SSR). A wide variety of cloning methods, host cells, and *in vitro* amplification methodologies are well-known to persons of skill.

As used herein, "nucleic acid sample" includes reference to a specimen suspected of comprising RG resistance genes. Such specimens are generally derived, directly or indirectly, from lettuce tissue.

The term "antibody" refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments or synthetic or recombinant analogues thereof which specifically bind and recognize analytes and antigens, such as a genus or subgenus of polypeptides of the invention, as described *supra*.

It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims.

EXAMPLES

The following examples are offered to illustrate, but not to limit the claimed invention.

Example 1

Example 1 describes the use of PCR to amplify RG genes from lettuce.

Multiple primers with low degeneracy, particularly at the 3' end, were designed based on the sequences of two known resistance genes from tobacco and flax.

DNA Templates

Lettuce genomic DNA was extracted from cultivar Diana and a mutant line derived from cultivar Diana using a standard CTAB protocol. To generate cDNA templates, RNA was isolated from cultivar Diana and the mutant following standard procedures; first strand cDNA was synthesized using Superscript reverse transcriptase from 1 Φg total RNA as specified by the manufacturer (Life Technologies). BAC (bacterial artificial chromosome) clones from the *Dm3* region were isolated from a BAC library of over 53,000 clones using marker AC15 that was known to be closely linked to *Dm3*. Bacterial plasmids containing clones of *L6* and *RPS2* were used as positive controls.

PCR with degenerate oligonucleotide primers

Oligonucleotide primers were designed based on conserved motifs in the nucleotide binding sites (NBS) of L6, RPS2, and N. Eight primers were made corresponding to the GVGKTT motif in the sense direction; each had 64-fold degeneracy. Six primers were made to the GLPLAL motif in the anti-sense direction; with either 16 or 256-fold degeneracy (Table 1).

Oligonucleotides included 14-mer adaptors of $(CUA)_4$ at the 5' end of the sense primers and $(CAU)_4$ at the 5' end of the antisense primers to allow rapid cloning of the PCR products into pAMP1 (Life Technologies).

PCR amplification was performed in 50 Φl reaction volume with 1 ΦM of each of a pair of sense and antisense primers. The templates were denatured by heating to 94EC for 2 min. This was followed by 35 cycles of 30 sec at 94EC, 1 min at 50EC, 2 min at 72EC, with a single final extension of 5 min at 72EC. 25 ng of genomic DNA or cDNA was used. BAC clones as templates required less. The final dNTP concentration was 0.2 mM; $MgCl_2$ was 1.5 mM.

Forty-eight combinations of sense and antisense primers were tested on a panel of nine templates consisting of two genomic DNA samples, two cDNA preparations, three BAC clones and plasmids containing *L6* and *RPS2* as positive controls. Amplification from *L6* and *RPS2* resulted in fragments of 516 and 513 repectively. Seven combinations of primers resulted in fragments of approximately this size with multiple templates (Table 2). Primers that gave RLG products were: PLOOPAA, PLOOPAG, PLOOPGA, PLOOPGG, PLOOPAC, GLPL3, GLPL4.

(Intentionally left blank)

Table 1

DEGENERATE PRIMER SEQUENCES for NBS PCR

Sense primers based on GVGKTT amino acid sequence from L6, N and rps2 PLOOP motif:

PLOOPAG 5' GGN GTN GGN AAA ACG AC 3'

PLOOPAA 5' GGN GTN GGN AAA ACA AC 3'

PLOOPAT 5' GGN GTN GGN AAA ACT AC 3'

PLOOPAC 5' GGN GTN GGN AAA ACC AC 3'

PLOOPGG 5' GGN GTN GGN AAG ACG AC 3'

PLOOPGA 5' GGN GTN GGN AAG ACA AC 3'

PLOOPGT 5' GGN GTN GGN AAG ACT AC 3'

PLOOPGC 5' GGN GTN GGN AAG ACC AC 3'

Antisense primers based on GLPLAL amino acid sequence:

GLPL1 5' AGN GCN AGN GGN AGG CC 3'

GLPL2 5' AGN GCN AGN GGN AGA CC 3'

GLPL3 5' AGN GCN AGN GGN AGT CC 3'

GLPL4 5' AGN GCN AGN GGN AGC CC 3'

GLPL5 5' AAN GCC AAN GGC AAA CC 3'

GLPL6 5' AAN GCC AAN GGC AAT CC 3'

TABLE 2. Characteristics of RLGs isolated from lettuce.

	Template	Primers	Number[a]	Size[b] (bp)	Copy number[c]	*Dm linkage*
RLG1	genomic DNA	PLOOPGA+GLPL6	6/6	522		*DM4*, *DM13*
	cDNA	PLOOPGA+GLPL6	1/5			
	genomic DNA	PLOOPAA+GLPL6	5/5			
	cDNA	PLOOPAA+GLPL6	1/1			
RLG2	BACH8	PLOOPGG+GLPL3	3/3	510		*DM1*, *Dm3*
RLG3	gemonic DNA	PLOOPGA+GLPL4	3/6	461		*Dm5* *Dm8*
RLG4	genomic DNA	PLOOPGA+GLPL4	1/6	524		

[a] Number of RLG sequences out of total number of clones sequenced.

[b] Size of fragment amplified from the nucleotide bindind domain.

[c] Estimated copy number from genomic Southern blot analysis and numbers of clones in the BAC library.

Example 2

Example 2 describes the genetic analysis used to obtain a preliminary indication of the linkage relationships of the amplified products and known clusters of resistance genes.

Bulked segregant analysis was performed to obtain a preliminary indication of the linkage relationships of the amplified products and known clusters of resistance genes. DNA from individuals were pooled for each susceptible and resistant bulk. Amplified products were then mapped by RFLP analysis from our intraspecific mapping population. Resistances from four clusters of resistance genes as well as over six hundred markers have now been mapped on this population. Linkage analysis was done using JIONMAP or MAPMAKER mapping programs. Due to a suppression of recombination in the *Dm3* region, sequences were mapped relative to *Dm3* using a panel of deletion mutants that provided greater genetic resolution than the mapping population (Anderson *et al.* 1996). All blots were washed twice at 63EC in 2x SSC/1% SDS for 20 min, followed by one wash at 63EC in 1x SSC/0.1% SDS for 10 or 30 min.

Most of the RLG sequences were analyzed by bulked segregant analysis (BSA) using pools of resistant and susceptible individuals for each of the four clusters of resistance genes. In genomic Southern analyses, all the RLGs revealed numerous fragments of varying intensity. The numbers of bands was highly dependent of the stringency of hybridization. BSA demonstrated that RLG1 was linked to the *Dm4,7* and *Dm13* clusters. Segregation analysis confirmed this linkage.

RLG2 was derived from BAC H8 that was known to be from the *Dm3* region. BSA with RLG2 demonstrated that the polymorphic bands that distinguished the parents of our mapping population mapped to the *Dm1,Dm3* cluster. Several bands absolutely cosegregated with *Dm1* or *Dm3*. To provide finer genetic resolution, RLG2 was also mapped using a panel of *Dm3* deletion mutants. A number of fragments were missing in largest deletion mutant demonstrating that several RLG2 family members are physically located very close to *Dm3*. No fragment was missing in all deletion mutants; however, this is not unexpected as there is extensive duplication within the region.

Example 3

Example 3 describes the screening of a bacterial artificial chromosome library.

Over 53,000 BAC clones containing lettuce genomic DNA were screened with two of the amplified products. High density filters each containing 1536 clones were hybridized to ^{32}P labelled probes. Filters were washed at 65EC with 40 mM Na_2PO_4/0.1% SDS for 5 min followed by 20 min in the same solution.

To isolate additional RLG sequences we screened our genomic BAC library. Clones were identified that hybridized to RLG1 and RLG2. Nearly all the clones that hybridized to RLG2 also hybridized to marker AC15 that had already been shown by deletion mutant analysis to be clustered around *Dm3*. This provided further evidence for clustering of RLG2 sequences.

Using primers conserved within each family, part of the NBS was amplified from each unique BAC clone and sequenced. This revealed that members within each family varied from 64% identical at the deduced amino acid level. The most divergent members only weakly cross-hybridized to each other. Currently, RLG sequences are

considered to be part of the same family of sequences if they are at least 55% identical at the deduced amino acid level and map to the same region of the chromosome.

Example 4:

Example 4 describes the cloning, identification, sequencing and characterization of RG polynucleotide sequences; including use of RG sequences from plasmid and PCR products.

Doubled stranded plasmid DNA clones and PCR products were sequenced using an ABI377 automated sequencer and fluorescently labelled di-deoxy terminators. Sequences were assembled using Sequencher (Genecodes), DNAStar (DNAStar) and Genetics Computer Group (GCG, Madison, WI) software. Database searches were performed using BLASTX and FASTA (GCG) algorithms.

Sequences flanking the NBS region for RLG2 and for some of RLG1 were obtained by a series of IPCR and the products sequenced directly. IPCR worked less well for RLG1. Therefore RLG1 was subcloned from a BAC clone into pBSK (Stratagene) and the double stranded plasmid sequenced by long range sequencing.

Initially, a total of 30 clones were sequenced. Three of these seven primer combinations yielded sequences that comprised continuous open reading frames with sequence identity to the NBS of known resistance genes. Seven out of 10 clones amplified from genomic DNA with the primer pair PLOOPGA/GLP6 were 522 bp long; they were identical to each other and named RLG1. All six clones amplified from genomic DNA or cDNA using the primers PLOOPAA/GLP6 were similar/the same as RLG1. All three clones sequenced from BAC clone H8 were 510 bp long, identical to each other but different from RLG1 and were therefore designated RLG2. The 11 clones sequenced from four other primer combinations had no similarity to any NBS motifs and therefore were not studied further. Therefore, sequencing resulted in the identification of clones containing NBS motifs representing four RLG sequences.

Comparison of the deduced amino acid sequences of RLG1 and RLG2 to those of known resistance genes revealed that RLG1 and RLG2 are as similar to each other as they are to resistance genes from other species and that this is the same level of identity shown between the known resistance genes (Table 3). The percent identity (upper quadrant) and percent identity (lower quadrant) were determined using the MEGALIGN

routine of the DNASTAR package. Identity refers to the proportion of identical amino acids; identity refers to the proportion of identical and similar amino acids and takes into account substitutions of amino acids with similar chemical characteristics. RG1 and RG2 are as similar to each other and to cloned resistance genes as cloned resistance genes from a variety of species are to each other. L6, resistance to *Melampsora lini* in flax (Lawrence *et al.*, 1995). *N*, resistance to tobacco mosaic virus in tobacco (Whitham *et al.*, 1994). *PRF*, required for resistance to *Pseudomonas syringae* in tomato. *RPS2*, resistance to *Pseudomonas syringae* in *Arabidopsis thaliana* (Bent *et al.*, 1994; Mindrinos *et al.*, 1994). RPM1, resistance to *Pseudomonas syringae* pv. *maculicola* in *A. thaliana* (Grant *et al.*, 1995). The initial RG1 and RG2, sequences were amplified from lettuce using degenerate primers.

Table 3

IDENTITIES OF RESISTANCE GENE HOMOLOGUES

		RG1	RG2	RG3	RG4	N gene	RPS2
Lettuce	RG1	***	22.7	15.0	29.2	25.4	23.8
Lettuce	RG2		***	32.2	21.6	22.7	33.0
Lettuce	RG3			***	17.2	15.0	32.8
Lettuce	RG4				***	44.3	22.7
Tobacco	N gene					***	21.6
Arabidopsis	RPS2						***

The regions homologous to the primers are included in this analysis as the genomic sequences for RLG1 and RLG2 were determined by IPCR. Interestingly, the genomic sequences for RLG1 exactly matched that of the primers used.

To obtain further evidence that we had amplified resistance genes, we amplified the regions flanking the NBSs of RLG1a and RLG2a by IPCR of BAC clones. These products were then directly sequenced without cloning to minimize the introduction of PCR artifacts. Sequence analysis of the 5' regions failed to detect any homology to known resistance genes. However, the sequence of the 3' region contained leucine-rich

repeats (LRRs). When this sequence was used to search GENBANK using BLASTX, it detected identity to the *Arabidopsis* resistance gene, *RPS2*. This region does not contain as regular LRRs as in some resistance genes; however, the repeat structure seems to be consistent with that of the flax resistance gene, *L6*. Therefore, the presence of an LRR region is further evidence that the sequences we amplified using degenerate oligonucleotide primers are probably resistance genes.

The sequences of the IPCR products also provided the genomic sequences of the regions complementary to the sequences of the degenerate oligonucleotide primers. The genomic sequences for RLG1 were identical to one of the primers in the mixture. The RLG sequences are resistance genes as supported by three criteria: the presence of multiple sequence motifs characteristic of resistance genes, genetic cosegregation with known resistance genes, and their existence as clustered multi-gene families. The presence of LRR regions in a similar position relative to the NBS as in cloned resistance genes provides stronger evidence than relying solely sequence similarity between NBS regions. The clustering of RLG sequences at the same position as the known clusters of resistance genes make them strong candidates for encoding resistance genes. The hybridization patterns and genetic distribution of the RLG sequences are similar to that of cloned resistance genes in other species. Most of these hybridize to small multigene families and preliminary genetic evidence indicates that they are clustered in the genome. Therefore, the degenerate primers that we designed from other resistance genes seemed to have been specific enough to amplify resistance genes rather than P-loop containing proteins in general.

(intentionally left blank)

SEQ ID NO: 1

RLG1A

[Strand]

1 ATCGTAACCGTTCGTACGAG ANCGCTGTCCCTCCTTCATC TTTTGTCATATGTCATATTC TCATNNATTNTGCCACATNT
81 AATTTTGTGGTTATTTTAAA TTAATTTTTATTCCACATGT CATTTTATGAGTTTTTCTAT TTTATTGAGTTTCACATAAT
161 ATTTAAATGTAATAACAATA AATGCATATTTATTTTTCTT TAAATAAACGCATATAATAT ATAGATTAAAATCATATAAT
241 ACATAGGTTAAACTCATATA ATACATATGTTCATCCCCAG TTTATTTATATGTCTCATCC TTAATTTATTTATTATTTAT
321 TTATTAGAGTAGATGATCTT TGTGATATTAAAAATTTAAT TTGTTCAAAATTTAAAATTA TTAATAATCCCACAATTTGA
401 ATAAAATTAAAAAAAATGGN CCCACCATTAGTCCATCACT TTTTCAGCTCATCAATATCG TGAGTATTCTCCTTCGTTTC
481 CACCCTAATCAATATTTCCA GCGAATGACAGACTCCTACG GCGTTTCTGAATTTGCGTTC CGACACTGTTCATTGAAGGA
561 GATAATAAATCAAATGGAGC TGCTCCAATGTTCATTGCTG ATGAAAGGTGAATTGTATGT GAAGANAATGTCAGCGATCN
641 ATCTCCATCCGGAACCCACC ACATTATCAGTGTACCACCA AACCACTCAAAACGGYGGAA GTAGRRAKACWRKAAAGTCA
721 TGAAGAATAGATTATTTTTG TCCTCATGGGCTGACTGAGG AGCGGGTTTAGTTCATCATT TTTCTTTGANCAAAGAATTA
801 TCGGTCCATCGAATTTTTAC ATCGACAAAGAAGTTTCACT TCGCAATGTTTTGTTAAACA ATTTTTAATCTTTTTATCTT
881 TTCGTTGAAACTCCTCAATT GCAACTTGCAACTTGCAACT TTTGGGCCCACAAATTTGTG GTGGGCGTTAATTTAATCCA
961 CATATTCACTGTAAACAATA ATTCAAATCGATCTCTGTTC ATCCAATTCATCAACATCTC TTGATAATTGAAATCATTCA
1041 CGCTTCATCCATTTCATCCA CATCTATACTATATTCTCTG CTCTTATCATATTAAACGAT GGCTGAAATCGTTCTTTCTG
1121 CCTTCTTGACAGTGGTGTTT GAAAAGCTGGCATYTGAAGC CTTGAAGAAGATTGTTCGCT CCAAAAGAATTGAATCTGAG
1201 CTTAAGAAATTGAAGGAGAC ATTAGACCAAATCCAAGATC TGCTTAACGATGCTTCCCAG AAGGAAGTAACTAATGAAGC
1281 CGTTAAAAGATGGCTGAATG ATCTCCAACATTTGGCTTAT GACATAGACGACCTACTTGA TGATYTTGCAACTGAAGCTG
1361 TTCAWCGTGAGTTGACCGAG GAGGGTGGAGCCTCCTCCAG TATGGTAAGAAAACTAATCC CAAGTTGTTGCACAAGTTTC
1441 TCACAAAGTAATAGGATGCA TGCCAAGTTAGATGATATTG CCACCAGGTTACAAGAACTG GTAGAGGCAAAAAATAATCT
1521 TGGTTTAAGTGTGATAACAT ATGAAAAGCCAAAAATTGAA AGGTATGAGGCGTCTTTGGT AGATGAAAGCGGTACTGTCG
1601 GACGTGAAGATGATAAGAAA AAATTGCTGGAGAAGCTGTT GGGGGATAAAGATGAATCAG GGAGTCAAAACTTCAGCATC
1681 GTGCCCATAGTTGGTATGGG TGGAGTTGGTAAAACAACTC TAGCTAGACTTTTGTATGAT GAAAAGAAAGTGAAGGATCA
1761 CTTCGAACTCAGGGCTTGGG TTTGTGTTTCTGATGAGTTC AGTGTTCCCAATATAAGCAG AGTTATTTATCAATCTGTGA
1841 CTGGGGAAAAGAAGGAGTTT GAAGACTTAAATCTGCTTCA AGAAGCTCTTAAAGAGAAAC TTAGGAACCAGCTATTTCTA
1921 ATAGTTTTGGATGATGTGTG GTCTGAAAGCTATGGTGATT GGGAGAAATTAGTGGGCCCA TTCCTTGCGGGGTCTCCTGG
2001 AAGTAGAATAATCATGACAA CTCGGAAGGAGCAATTGCTC AGAAAGCTGGGCTTTTCTCA TCAAGACCCTCTGGAGGGTC
2081 TATCACAAGATGATGCTTTG TCTTTGTTTGCTCAACACGC ATTTGGTGTACCAAACTTTG ATTCACATCCAACACTAAGG
2161 CCACATGGAGAACTGTTTGT GAAGAAATGTGATGGCTTAC CTCTAGCYTTAAGAACACTT GGAAGGTTATTAAGGACAAA
2241 AACAGACGAGGAACAATGGA AGGAGCTGTTGGATAGTGAG ATATGGAGGTTAGGAAAGAG CGATGAGATTGTTCCGGCTC
2321 TTAGACTAAGCTACAATGAT CTTTCTGCCWCTTTGAAGCT RTTRTTTGCATAYTGCTCCT TGTTTCCCAAGGACTATGAG
2401 TTTGACAAGGAGGAGTTGAT TCTATTGTGGATGGCAGAAG GGTTTTTGCACCAACCAACT AYAAACAAGTCAAAGCAACG
2481 KTTGGGTCTTGAATATTTTR AAGAGTTRTTGTCAAGRTCR TTTTTTCAACATGCTCCTAA TRRCAAATCSTTGTTTGTGA
2561 TGCATGACCTAATGAATGAT TTGGCTACATTTGTTGCTGG AGAATTTTTTTCAAGGTTAG ACATAGAGATGAAGAAGGAA
2641 TTTAGGATGSAATCTTTGGA RAAGCACCGMCATATGTCAT TTGTATGTGAGRATTACATA GGTTACAAAARGTTCGAGCC
2721 ATTTAGAGGAGCTAAAAATT TGAGAACATTTTTAGCATTG TCTGTTGGGGTGGTAGAAGA TTGGAAGATGTTTTACTTAT
2801 CAAACAAGGTCTTGAATGAC WTACTTCARGATTTACCATT GTTAAGGGTCCTRAKTTTGA TTRRTCTTAYAATAASYRAG
2881 GTACCARAAKTCGTSGGTAG TATGAASCACTTGCGGTATC TTAATCTATCWGRAACTTWA ATCACMCATTTACCGGAAWA
2961 TKTCTGCAATCTTTATAATT TACARACCCTGATTGTKTCT GGCTGTGAMTATTTAGTTAA KTTGCCCAARACCTTCTCAA
3041 ASCTTAAAAATTTGCASCAT TTTGACATGAGGGRTACTCC KAAKTTRAARAACATGCCCT TARGGATTGGTGARTTGAAA
3121 ARTCTACAAACTCTCTTYMG TAACATTGGCATAGCAATAA CCGAGCTTAAGAACTTGCAM AAYCTCCATGGGAAARTTTG
3201 TATTGGCGGGCTGGAAAAAA TGGAAAATGCMGTKGGATGC ACGTTAAGCGAACTTGTCTC AAAAAAGGTTWAATGARTTA
3281 NAAACTGGRWTKGGGGGTGA TRAATTTAATGTTTTCCGAA ATGGGAACACTTGAAAAAGA AGTCCTCAATGAAGTGATGC
3361 CTCATAATGGTACTCTANAA AAAACCCANAATTATGTCTA TAGGGGGTATAGAGTTTCCA AATTGGGTTGGTTNCACTAA
3441 GGGTTTCTGAAACTAGAGAT GTGTTCATGGTGTATGAAAA AGANTGTTTTACGTAGTTTC ATCAATCACCAAGTGGGAAA
3521 TAGATGATATTTTCAGGGCY TACTGATGAGATGTGGAGAG GTATGATAGGGTNTCTTGGG GCGGTAGAAGAAATAAGCAT
3601 CCATTCTTGTAATGAAATAA GATATYTGTGGAATCAGAA GCAGAGGCAAGTAAGGTTCT TATGAATTTAAAGAAGTTGG
3681 ATTTAGGTGAATGTGAAAAT TTGGTGAGTTTAGGGGAGAA AAAGGAGGATAATCATAATA TTAATAGTGGGAGCAGCCTA
3761 ACATCTTTTAGGAGGTTGAA TGTATGGAGATGTAACAGCT TGGAGCATTGCAGGTGTCCA GATAGCATGGAGAATTTGTA
3841 TATGCACATGTGTGATTCAA TNACATCCGTCTCCTTCCCA ACAGGAGGAGGACAGAAGAT CAAGTCACTTACCATCACTG
3921 ATTGCAAGAAGCTTTCGGAA GAGGAGTTGGGAGGACGAGA GAGGACAAGAGTGCTTATAA ACTCAAAAATGCAGATGCTT
4001 GAATCAGTAGATATACGTAA TTGGCCAAATCTGAAATCTA TCAGTGAATTGAGTTGCTTC ATTCACCTGAACAGATTATA
4081 TATATCAAACTGTCCGAGTR TGGAGTCATTTCCTGACCAT GAGTTGCCAAATCTCACCTC CTTAACAGATCGAAGGAGAG
4161 GACAGCGATTTTCGTACGAA CGGTTACGATTCGACTGGCC GTCGTTTT

SEQ ID NO:2

RLG1B

[Strand]

```
1     AACCGTTCGT ACGAGAATCG CTGTCCTCTC CTTCCTGTAA TATAATGATA AGAAAAAATA TGATTAAAGG
71    TTTAAATCCA AAATCCATTA TTCCACCGGT GATATGATGC ACTAGCTGTA GTATGCAAAA ACAGTATTAT
141   AAATGCTAAC CAAAACAGCA GCTAAGAAAC AATATAAATA ATGGTTTGAA TCGTCCTTTC TCCGTACACT
211   CATTTCTTCC AAATCCCTAT CATTCATACA TACAAGTGCT CCCATATTAG GTTTTCACTA TAAGCAATGG
281   CTGAAATCCT TGGTTCTGCG TTCTTTGCGG TGTTCTTTGA AAAGCTTGCT TCTGAAGCCT TGAAGAGGGT
351   TGCTTGCTCC AAAGTAATTG ACAAGGAGCT CGAGAAATTG AATAGCTCAT GAATCAATAT AAAAGCTCTG
421   CTCAATGATG CTTCTCAGAA GGAAATAAGT AAGGAAGCTG TTAAAGAATG GTTGAATGCT CTTCAACATT
491   TGCCTTACGA CATAGATGAT CTACTTGGCG ATTTGGCAAC CAAAGCTATC CATCGTAAGT TCTCTGAGGA
561   ATACGGGGCC ACCATCAACA AGGTACGAAA GTTAATTCCA TCTTGTTTCT CTAGTTTGTC AAGTACTAAG
631   ATGCGCAACA AGATACATAA TATTACCAGC AAGTTACAAG AACTATTAGA AGAGAGAAAT AATCTTGGAT
701   TATGTGAAAT TGGTGAAAGC CGAAAACTTC GAAATAGAAA ATCAGAGACC TCTNTGCTAG ATCCATCTAG
771   TATTGTTGGA CGCACAGATG ATAAGGAAGC GTTGCTTCTC AAGCTATATG AACCATGTGA TAGAAACTTT
841   AGCATCTTGC CNATAGTTGG TATGGGTGGG TTAGATAAGA CCACTTTAGG TAGACTTTTG TATGATNAAA
911   TGCAAGTGAA GGATCACTTC GAACTCAAGG CGTGGGTTTG TGTTTCTGAT GAGTTTGATA TCTTCGGTAT
981.  AAGCAAAACC ATTTTCGAAT CGATAGAGGG GGGAAACCAA GAGTTTAAGG ATTTAAATCT GCTTCAGGTG
1051  GCTTTAAAGG AGAAAATCTC AAAGAAACGA TTTCTTGTTG TTCTTGATGA TGTATGGAGC GAGAGCTATA
1121  CTGATTGGGA AATTCTAGAA CGTCCATTTC TAGCAGGAGC ACCAGGAAGT AAAGTAATCA TCACAACCCG
1191  CAAGTTGTCG TTGCTAAACC AATTGGGTCA TGATCAACCA TACCAATTGT CTGATTTGTC ACATGACAAT
1261  GCTCTATCCT TATTTTGTCA ACACGCATTT GGTGTAAATA GCTTTGATTC ACATCCGATA CTTAAACCAC
1331  ATGGTGAAGG TATTGTTGAA AAATGTGATG GTTTGCCATT GGCTTTGATT GCACTTGGGA GGTTATTGAG
1401  GACAAAAAGA GATGAGGAAG AATGGAAGGA ACTATTGAAT AGTGAGATAT GGAGGTTAGG AAAGAGAGAT
1471  GAGATTATTC CGGYTCTTAG ACTAAGCTAT AATGATCTTT CTGCCTCTTT GAAGCAGTTG TTTGCATATT
1541  GCTCCTTGTT CCCCAAAGAC TATGTGTTCA ACAAGGAGAA GTTGATTTTA TTATGGATGG CAGAAGGGTT
1611  TTTGCACAAT GAAAATACAA ACAAGTCAAT GGAACGCTTA GNTCTTGAAT ATTTTGACGA CTTGTTGTCA
1681  AGGTCATTTT TTCAACATGC ACTCGATGAC AAATCGTTGT TTGTGGTGCA CGACCTCATG AATGACTTGG
1751  CCACATCTGT TGCTGGAGAT TATTTTTTAA GATTAGACAT TGAAATGAAA AAGGAAGCTT TGGAAAAATA
1821  CCGACATATG TCATTTGTTT GTGAGAGTTA CATGGTTTAC AAAAGGTTCG AACCATTTAA AGGAGCTAAA
1891  AAATTGAGAA CTTTCTTAGC AATGCCTGTT GGGATGATAA AAAGTTGGAC AACATTTTAC TTATCAAATA
1961  AGGTCCTTGA TGACTTACTT CACGAATTAC CATTGTTGAG AGTTCTAAGT TTGAGTTATC TTAGCATCAA
2031  GGAGGTACCT GAAATAATAG GCAATTTGAA ACACTTGCGG TATCTTAATT TATCACACAC GAGTATCACA
2101  CATTTACCAG AAAATGTCTG CAATCTTTAC AACTTACAAA CATTGATCCT TTGTGGCTGT TGTTTTATAA
2171  CCAAGTTTCC CAACAACTTC TTAAAGCTTA GAAATTTACG GCATTTGGAC ATTAGCGATA CTCCCGGTTT
2241  GAAGAAGATG TCCTCGGGGA TTGGTGAATT GAAGAACCTA CACACYCTCT CCAAGCTCAT TATTGGAGGT
2311  GAAAATAGAC TAAACGAGCT TAAGAACTTA CAAAATCTCC ATG
```

RLG1b - Diana
[Strand]

```
1    TACTACTACT AGAATTCGGT GTTGGTAAGA CGAcTCTAGC TAGACTTTTG TATGAGGAAA TGCAAGGGAA
71   GGATCACTTC GAACTTAAGG CGTGGGTATG TGTTTCTGAT GAGTTTGATA TCTTCAATAT AAGCAAAATT
141  ATCTTACAAT CGATAGGTGG TGGAAACCAA GAATTTACGG ACTTAAACCT GCTTCGAGTA GCTTTAAAAG
211  AGAAGATcTC AAAGAAAAGa TTTCTTCTTG TTCTTGATGA TGTTTGGAGT GAAAGCTATA CCGATTGGGA
281  AATTNTAGAA CGCCCATTTC TTGCAGGGGC ACCTGGAAGT AAGATTATTA TCACCACCCG GAAGCTGTCA
351  TTGTTAAACA AACTCGGTTA CAATCAACCT TACAACCTTT CGGTTTTGTC ACATGAGAAT GCTTTGTCTT
421  TATTCTGTCA GCATGCATTG GGTGAAGATA ACTTCAATTC ACATCCAACA CTTAAACCAC ATGGCGnAGG
491  TATTGTTGAA AAATGTGATG GaTTGCCATT GGCATTGTCG ACATGATGAT GATG
```

SEQ ID 137

SEQ ID NO:3

RLG1C

[Strand]

```
1    TCCCGTGCAA CGTNTATCAT TCAGAAGNGC CCAAAGACCA NAGATNTGTT TAANGNTGNT TNTCAGAAGG
71   AAGTAATTGA TGAAGCTGTN AAAAGATGGC TGATTGATNT CCAACAATTG GCTTACGACA CTGANGACNA
141  ACTTGATGAT NTCGCAACAG AAGCTATTCA TCGTGAGTTG ATCCGTGAAA CTGGAGCTTC CNCCAGCATG
211  GTAAGAAAGC TAATCCCAAG TTGTTGCACA AGTTTCTCAC AAAGTAATAG GATGCATGCC AGGTTAGATG
281  ATATTGCCGC TAAGTNACAA GAACTGGTAG AGGCGAAAAA TAATCTTGGT TTAAGTGTGA TAACATACGA
351  AAAACCCAAA ATTGAAAGAG ATGAGGCGTN TTTGGTAGAT GCAAGTGGTA TCATTGGACG TGAAGATGAT
421  AAGAAAAAAT TGCTTCAGAA GCTGTTGGGG GATACTTATG AATCAAGTAG TCAAAACTTC AACATCGTGC
491  CCATAGTTGG TATGGGTGGG GTAGGTAAAA CAACTCTAGC TAGACTTTTG TATGATGAAA AAAAAGTGAA
561  GGATCACTTC GAACTCAGGG TTTGGGTTTG TGTTTCTGAT GAGTTCAGTG TTCCCAATAT AAGCAGAGTT
631  ATCTATCAAT CTGTGACTGG TGAAAACAAA GAATTTGCAG ATTTAAATCT GCTTCAAGAA GCCCTTAAAG
701  AGAAACTTCA GAACAAACTA TTTCTAATAG TTTTAGATGA TGTATGGTCT GAAAGCTATG GTGATTGGGA
771  GAAATTAGTG GGCCCATTTC ATGCTGGGAC TTCTGGAAGT AGAATAATCA TGACTACTCG GAAGGAGCAA
841  TTACTCAAAC AGCTGGGTTT TTCTCATGAA GACCCTCTGC ATAGTATAGA CTCCCTGCAA CGTCTATCAC
911  AAGAAGATGC TTTGTCTTTG TTTTCTCAAC ACGCATTTGG TGTACCTAAC TTTGATTCAC ATCCAACACT
981  AAGGCCATAT GGGGAACAGT TTGTGAAAAA ATGTGGGGGA TTGCCTTTGG CCTTGT
```

SEQ ID NO:4

RLG1D
[Strand]

```
1     CNTACCNTTC TACGAGATCG CTGTCCCTCC TCGATCTGCT TAACGATGCT TCCCAGAAGG AAGTNACTAA
71    TGAAGCCGTT AAAAGATGGC TGAATGATCT CCAACATTTG GCTTATGACA TANACGACCT ACTTGATGAT
141   CTTGCAACAS AAAGCTATTC NTCSTGAGTT GACCGANGAA GGTGGAGCCT CCACCAGTAT GGTAAGAAAA
211   CTAATCCCAA GTTGTTGCAC AAGTTTCTCA CAAAGTTATA GGATGCATGC CAAGTTAGAT GATATTGCCA
281   CCAGGTTACA AGAACTGGTA GAGGCAAAAA ATAATCTTGG TTTAAGTGTG ATAACATATG AAAAGCCCAA
351   AATTGAAAGG TATGAGGCAT CTTTGGTAGA CGAAAGTGGT ATTTTTGGAC GTTNAGATGA TNAGAAAAAA
421   TTGATGGAGA AGCTGTTGGA GGATAAAGAT GAATCCGGAG TCNAAACTTC AGCATCCTGC CCATAATTGG
491   TATGGGTGGA GTTGGCNAAA CAACTCTAGC TAGACTCTTG TTTGATGAAA AGACAGTGAA GGATCACTTC
561   GAACTCAGGG CTTGGGTTTG TGTTTCTGAT GAATTCAGTA TTCTCAACAT AAGCAAAGTT ATCTATCAAT
631   CTGTGACCGG GGAAAAGAAA GAGTTTGAAG ACTTAAATCT GCTTCAAGAA GCTCTTAGAG GGAAACTACA
701   AAACAAACTA TTTCTAATAG TTTTGGATGA TGTATGGTCG GAAAGCTATG GTGATTGGGA GAAATTAGTG
771   GGCCCATTTC ATGCTGGGAC TTCTGGAAGT AGAATAATCA TGACTACTCG GAAGGAGCAA TTACTCAAAC
841   AGTTGGGTTT TTCTCATCAA GACCCTCTGC GTTGTATAGA CTCCCTGCAA CGTCTATCAC AAGATGATGC
911   TTTGTCTTTG TTTGCTCAAC ACGCATTTGG TGWCCA
```

RLG1E
[Strand]

```
1    TCTAGCTAGA CTTTTGTATG ACGAGATGCA AGAGAAGGAT CACTTCGAAC TCAAGGCGTG GGTTTGTGTT
71   TCTGATGAGT TTGATATATT CAATATAAGC AAAATTATTT TCCAATCGAT AGGAGGTGGA AACCAAGAAT
141  TTAAGGACTT AAATCTCCTT CAAGTAGCTG TAAAAGAGAA GATTTCAAAG AAACGATTTC TACTTGTTCT
211  TGATGATGTT TGGAGTGAAA GCTATGCGGA TTGGGAAATT CTGGAACGCC CATTTCTTGC AGGGGCAGCC
281  GGAAGTAAAA TTATCATGAC GACCCGGAAG CAGTCATTGC TAACCAAACT CGGTTACAAG CAACCTTACA
351  ACCTTTCCGT TTTGTCACAT GACAGTGCTC TCTCTTTATT CTGTCAGCAT GCATTGGGTG AAGATAACTT
421  CGATTCACAT CCAACACTTA AACCACATGG CGAAGGCATT GTTGAAAAAT GTGCT
```

SEQ ID NO:5

RLG1F

[Strand]

1 ATTTTCNGCT CNAAACAAAN AAAAGCAATG GCTGAAATCT TTCTTTCNGC ATTCTAGACC AGTATTCTTT
71 GAAAAGNTGG CTTCTGAAGC CTTGAAGAAG ATCGCTCGCT TCCATCGGAT TGATTCTGAG CTCAAGAAAC
141 TGAAGAGGTC ATTAATCCAG ATCAGATCTG TGCTTAATGA TGCTTCTGAG AAGGAAATAA GTGATGAAGC
211 TGTTAAAGAA TGGCTGAATG GTCTCCAACA TTTGTCTTAC GACATAGACG ACCTACTTGA TGATTTGGCA
281 ACCGAAACTA TGCATCGTGA GTTGACCCAC GGATCTGGAG CCTCCACCAG CTTGTAAGAA AGATAATCCC
351 AACTTGTTGC ACAGATTTCT CACTAAGTAG TAAGATGCGT AACAAGTTAG ATAATATTAC CATCAAGTTA
421 CAAGAACTGG TAGAGGAAAA AGATAATCTT GGCTTAAGTG TGAAAGGTGA AAGCCCAAAA CATACCAACA
491 GAAGATTACA GACCTCTTTG GTAGATGCAT CTAGCATTAT TGGTCGTGAA GGTGATAAGG ATGCATTGCT
561 CCATAAGCTG CTGGAGGATG AACCAAGTGA TAGAAACTTT AGCATCGTGC CAATAGTTGG TATGGGTGGT
631 GTGGGTAAGA CGACTCTAGC TAGACTTTTG TATGACGAGA TGCAAGAGAA GGATCACTTC GAACTCAAGG
701 CGTGGGTTTG TGTTTCTGAT GAGTTTGATA TCTTCAATAT AAGCAAAGTT ATCTTCCAAT CGATAGGTGG
771 TGGARACCAA GAATTTAAGG ACTTAAATCT CCTTCAAGTA GCTGTAAAAG AGAAGATTTC AAAGAAACGA
841 TTTCTWYTTG TTCTGGATGA TGTTTGGAGT GAAAGCTATA CAGAATGGGA AATTCTAGCA CGTCCATTTC
911 TTGCAGGGGC ACCAGGAAGT AAGATTATCA TGACGACCCG GAAGTTGTCG TTGCTAACCA AACTCGGTTA
981 CAATCAACCT TACAACCTTT CSGTTTTGTC ACATGATAAT GCTYTGTCTT TATTCTGTCA GCAYGCATTG
1051 GGTGAAGATA ACTTCGATTC ACATCCAACA CTTAAACCAC ASGGTGAAAG TATTGTTGAA AAATGTGACG
1121 GTTTACCATT GGCTTTRATT GCACTTGGGA GRTTGTTGAR GACAAAAACA GATGAGGAAG AATGGAARGA
1191 AGTGTTGAAT AGTGAAATAT GGGGGTCAGG AAAGGGAGAT GAGATTGTTC CGGCTCTTAA ACTAAGCTAC
1261 AATGATCTCT CTGCCTCTTT GAAGAAGTTG TTTGCATACT GCTCCTTGTT CCCAAAAGAC TATGTGTTCG
1331 ATAAGGAGGA GTTGATTTTG TTGTGGATGG CAGAAGGGTT TTTGCACCAA TCAACCACAA GCAAGTCBAT
1401 GGAACGCTTG GGHCATGAAG GTTTTGATGA ATTGTTGTCA AGATCATTTT TTCAACATGC CCCTGATGCC
1471 AAATCGATGT TTGTGATGCA TGACCTGATG AATGACTTGG CHACATCTGT TGCTGGAGAT TTTTTTTCAA
1541 GGATGGACAT TGAGATGAAG AARGAATTTA GGAAGGAAGC TTTGSAAAAG YAYCGCCATA TGTCAWTTGT
1611 TTGTGAKGAT TACATGGTKK ACAAAAGGTT CRAGCCATTS ACAAGGAGCT AG

SEQ ID NO: 6

RLG1G
[Strand]

1 GTGAAGGATC ACTTCGAACT CAGGGCTTGG GTTTGTGTTT CTGATGAATT TAATATCCTC AATATAAGCA
71 AAGTAATTTA TCAATCTGTA ACCGGGGAAA AAAAGGAGTT TGAAGACTTA AATCTGCTTC AAGAAGCTCT
141 TAAAGAAAAA CTTTGGAATC AGTTATTTCT AATAGTTCTG GATGATGTGT GGTCTGAAAG CTATCGTGAT
211 TGGGAGAAAT TAGTGGGCCC ATTTTTTTCG GGGTCTCCTG GAAGTATGAT TATCATGACA ACTCGGAAGG
281 AGCAATTGCC AAGAAAGCTG GGTTTTCCTC ATCAAGACCC TTTGCAAGGT CTATCACATG ACGATGCTTT
351 GTCTTTGTTT GCTCAACACG CATTTGGTGT ACCA

SEQ ID NO:7

RLG1H
[Strand]

```
1    TCTAGCTAGA CTTTTGTATG AGGAAATGCA AGGGAAGGAT CACTTCGAAC TCAAGGCGTG GGTATGTGTT
71   TCTGATGAGT TTGATATCTT CAATATAAGC AAAATTATCT TACAATCGAT AGGTGGTGGA AACCAAGAAT
141  TTACGGACTT AAACCTGCTT CAAGTAGCTT TAAAAGAGAA GATCTCAAAG AAAAGATTTC TTCTTGTTCT
211  TGATGATGTT TGGAGTGAAA GCTATACCGA TTGGGAAATT CTAGAACGCC CATTTCTTGC AGGGGCACCT
281  GGAAGTAAGA TTATTATCAC CACCCGGAAG CTGTCATTGT TAAACAAACT CGGTTACAAT CAACCTTACA
351  ACCTTTCGGT TTTGTCACAT GAGAATGCTT TGTCTTTATT CTGTCAGCAT GCATTGGGTG AAGATAACTT
421  CAATTCACAT CCAACACTTA AACCACATGG CGAAGGTATT GTTGAAAAAT GTGAT
```

SEQ ID NO:8

RLG1I
[Strand]

1	TCTAGCTAGA	CTTGTGTATG	ATGAGATGCA	AGAGAAGGAT	CACTTTGAAC	TCAAGGCGTG	GGTATGTGTT
71	TCTGATGAGT	TTGATATATT	CAATATAAGC	AAAATTATTT	TCCAATCGAT	AGGAGGTGGA	AACCAAGAAT
141	TTAAGGACTT	AAACCTCCTT	CAAGTAGCTG	TAAAAGAGAA	GATTTTAAAG	AAACGATTTC	TTCTTGTTCT
211	TGACGACGTT	TGGAGTGAAA	GCTATGCCGA	TTGGGAAATT	NTGGAACGCC	CATTTCTTGC	AGGGGCAGCC
281	GGAAGTAAAA	TTATCATGAC	AACCCGAAAG	CAGTCATTGC	TAACCAAACT	CGGTTACAAG	CAACCTTACA
351	ACCTTTCCGT	TTTGTCACAT	GACAGTGCTC	TGTCTTTATT	CTGTCAGCAT	GCATTGGGTG	AAGGTAACTT
421	CGATTCACAT	CCAACACTTA	AACCACATGG	CGAAGGCATT	GTTGAAAAAT	GTGCTGGATT	GCCATTGGCA
491	TTGTCGACA						

SEQ ID NO. 9

RLG1J
[Strand]

```
1    TACTACTACT AGAATTCGGT GTTGGTAAGA CGAcTCTAGC TAGACTTTTG TATGAGGAAA TGCAAGGGAA
71   GGATCACTTC GAACTTAAGG CGTGGGTATG TGTTTCTGAT GAGTTTGATA TCTTCAATAT AAGCAAAATT
141  ATCTTACAAT CGATAGGTGG TGGAAACCAA GAATTTACGG ACTTAAACCT GCTTCGAGTA GCTTTAAAAG
211  AGAAGATcTC AAAGAAAAGa TTTCTTCTTG TTCTTGATGA TGTTTGGAGT GAAAGCTATA CCGATTGGGA
281  AATTNTAGAA CGCCCATTTC TTGCAGGGGC ACCTGGAAGT AAGATTATTA TCACCACCCG GAAGCTGTCA
351  TTGTTAAACA AACTCGGTTA CAATCAACCT TACAACCTTT CGGTTTTGTC ACATGAGAAT GCTTTGTCTT
421  TATTCTGTCA GCATGCATTG GGTGAAGATA ACTTCAATTC ACATCCAACA CTTAAACCAC ATGGCGnAGG
491  TATTGTTGAA AAATGTGATG GaTTGCCATT GGCATTGTCG ACATGATGAT GATG
```

SEQ ID NO:10

RLG1A a.a.

IVTVRTR?LSLLHLLSYVIFS?I?PH?ILWLF.INFYSTCHFMSFSILLSFT.YLNVITINAYLFFFK.THIIYR
LKSYNT.VKLI.YICSSPVYLYVSSLIYLLFIY.SR.SL.Y.KFNLFKI.NY..SHNLNKIKKNGPTISPSLFQLINIV
SILLRFHPNQYFQRMTDSYGVSEFAFRHCSLKEIINQMELLQCSLLMKGELYVK?MSAI?LHPEPTTLSV
YHQTTQNGGSR?T?KS.RIDYFCPHGLTEERV.FIIFL.?KNYRSIEFLHRQRSFTSQCFVKQFLIFLSFR.NS
SIATCNLQLLGPQICGGR.FNPHIHCKQ.FKSISVHPIHQHLLIIEIIHASSISSTSILYSLLLSY.TMAEIVLS
AFLTVVFEKLA?EALKKIVRSKRIESELKKLKETLDQIQDLLNDASQKEVTNEAVKRWLNDLQHLAYDID
DLLDD?ATEAV?RELTEEGGASSSMVRKLIPSCCTSFSQSNRMHAKLDDIATRLQELVEAKNNLGLSVI
TYEKPKIERYEASLVDESGTVGREDDKKKLLEKLLGDKDESGSQNFSIVPIVGMGGVGKTTLARLLYDEK
KVKDHFELRAWVCVSDEFSVPNISRVIYQSVTGEKKEFEDLNLLQEALKEKLRNQLFLIVLDDVWSESY
GDWEKLVGPFLAGSPGSRIIMTTRKEQLLRKLGFSHQDPLEGLSQDDALSLFAQHAFGVPNFDSHPTLR
PHGELFVKKCDGLPLALRTLGRLLRTKTDEEQWKELLDSEIWRLGKSDEIVPALRLSYNDLSA?LKLLFA
YCSLFPKDYEFDKEELILLWMAEGFLHQPT?NKSKQRLGLEYF?ELLSRSFFQHAPN?KSLFVMHDLMND
LATFVAGEFFSRLDIEMKKEFRM?SLEKHRHMSFVCE?YIGYK?FEPFRGAKNLRTFLALSVGVVEDWK
MFYLSNKVLND?LQDLPLLRVL?LI?L?I??VP??VGSM?HLRYLNLS?T?ITHLPE??CNLYNLQTLIV
SGC?YLV?LPKTFS?LKNL?HFDMR?TP?LKNMPL?IGELK?LQTLF?NIGIAITELKNL?NLHGK?CIGG
LGKMENAVGCTLSELVSKKV?.??NW??G..I.CFPKWEHLKKKSSMK.CLIMVL?KKP?IMSIGGIEFPN
WVGSLRVSETRDVFMVYEK?CFT.FHQSPSGK.MIFSG?TDEMWRGMIG?LGAVEEISIHSCNEIRYLWE
SEAEASKVLMNLKKLDLGECENLVSLGEKKEDNHNINSGSSLTSFRRLNVWRCNSLEHCRCPDSMENLY
MHMCDS?TSVSFPTGGGQKIKSLTITDCKKLSEEELGGRERTRVLINSKMQMLESVDIRNWPNLKSISEL
SCFIHLNRLYISNCPS?ESFPDHELPNLTSLTDRRRGQRFSYERLRFDWPSF

SEQ ID NO:11

RLGIB a.a.

NRSYENRCPLLPVI...EKI.LKV.IQNPLFHR.YDALAVVCKNSIINANQNSS.ETI.IMV.IVLSPYTHFFQIPII
HTYKCSHIRFSLAMAEILGSAFFAVFFEKLASEALKRVACSKVIDKELEKLNSS.INIKALLNDASQKEIS
KEAVKEWLNALQHLPYDIDDLLGDLATKAIHRKFSEEYGATINKVRKLIPSCFSSLSSTKMRNKIHNITS
KLQELLEERNNLGLCEIGESRKLRNRKSETS?LDPSSIVGRTDDKEALLLKLYEPCDRNFSILPIVGMGGL
DKTTLGRLLYD?MQVKDHFELKAWVCVSDEFDIFGISKTIFESIEGGNQEFKDLNLLQVALKEKISKKRFL
VVLDDVWSESYTDWEILERPFLAGAPGSKVIITTRKLSLLNQLGHDQPYQLSDLSHDNALSLFCQHAFG
VNSFDSHPILKPHGEGIVEKCDGLPLALIALGRLLRTKRDEEEWKELLNSEIWRLGKRDEIIP?LRLSYND
LSASLKQLFAYCSLFPKDYVFNKEKLILLWMAEGFLHNENTNKSMERL?LEYFDDLLSRSFFQHALDDKS
LFVVHDLMNDLATSVAGDYFLRLDIEMKKEALEKYRHMSFVCESYMVYKRFEPFKGAKKLRTFLAMPV
GMIKSWTTFYLSNKVLDDLLHELPLLRVLSLSYLSIKEVPEIIGNLKHLRYLNLSHTSITHLPENVCNLYN
LQTLILCGCCFITKFPNNFLKLRNLRHLDISDTPGLKKMSSGIGELKNLHTLSKLIIGGENRLNELKNLQNL
H

SEQ ID NO:12

RLG1c a.a.

SRAT?IIQK?PKT?D?F????QKEVIDEAVKRWLID?QQLAYDT?D?LDD?ATEAIHRELIRETGAS?S
MVRKLIPSCCTSFSQSNRMHARLDDIAAK?QELVEAKNNLGLSVITYEKPKIERDEA?LVDASGIIGRED
DKKKLLQKLLGDTYESSSQNFNIVPIVGMGGVGKTTLARLLYDEKKVKDHFELRVWVCVSDEFSVPNIS
RVIYQSVTGENKEFADLNLLQEALKEKLQNKLFLIVLDDVWSESYGDWEKLVGPFHAGTSGSRIIMTTR
KEQLLKQLGFSHEDPLHSIDSLQRLSQEDALSLFSQHAFGVPNFDSHPTLRPYGEQFVKKCGGLPLAL

SEQ ID NO:13

RLG1D

?T?LRDRCPSSICLTMLPRRK?LMKPLKDG.MISNIWLMT?TTYLMILQ?KAI??ELT?EGGASTSMVRK
LIPSCCTSFSQSYRMHAKLDDIATRLQELVEAKNNLGLSVITYEKPKIERYEASLVDESGIFGR?DD?KK
LMEKLLEDKDESGVKLQHLPIIGMGGVG?TTLARLLFDEKTVKDHFELRAWVCVSDEFSILNISKVIYQS
VTGEKKEFEDLNLLQEALRGKLQNKLFLIVLDDVWSESYGDWEKLVGPFHAGTSGSRIIMTTRKEQLLK
QLGFSHQDPLRCIDSLQRLSQDDALSLFAQHAFG?

SEQ ID NO:14

RLGIE

LARLLYDEMQEKDHFELKAWVCVSDEFDIFNISKIIFQSIGGGNQEFKDLNLLQVAVKEKISKKRFLLVLD
DVWSESYADWEILERPFLAGAAGSKIIMTTRKQSLLTKLGYKQPYNLSVLSHDSALSLFCQHALGEDNF
DSHPTLKPHGEGIVEKCA

SEQ ID NO:15

RLGIF

FSA?NK?KQWLKSFF?HSRPVFFEK?ASEALKKIARFHRIDSELKKLKRSLIQIRSVLNDASEKEISDEA
VKEWLNGLQHLSYDIDDLLDDLATETMHRELTTDLEPPPACKKDNPTCCTDFSLSSKMRNKLDNITIKL
QELVEEKDNLGLSVKGESPKHTNRRLQTSLVDASSIIGREGDKDALLHKLLEDEPSDRNFSIVPIVGMGG
VGKTTLARLLYDEMQEKDHFELKAWVCVSDEFDIFNISKVIFQSIGGG?QEFKDLNLLQVAVKEKISKKR
FL?VLDDVWSESYTEWEILARPFLAGAPGSKIIMTTRKLSLLTKLGYNQPYNLSVLSHDNALSLFCQHA
LGEDNFDSHPTLKP?GESIVEKCDGLPLALIALGRLL?TKTDEEEWKEVLNSEIWGSGKGDEIVPALKLS
YNDLSASLKKLFAYCSLFPKDYVFDKEELILLWMAEGFLHQSTTSKSMERLGHEGFDELLSRSFFQHAPD
AKSMFVMHDLMNDLATSVAGDFFSRMDIEMKKEFRKEAL?K?RHMS?VC?DYMV?KRF?P?TRS.

SEQ ID NO:16

RLG1 G

VKDHFELRAWVCVSDEFNILNISKVIYQSVTGEKKEFEDLNLLQEALKEKLWNQLFLIVLDDVWSESYR
DWEKLVGPFFSGSPGSMIIMTTRKEQLPRKLGFPHQDPLQGLSHDDALSLFAQHAFGVP

SEQ ID NO:17

RLG 1 H

LARLLYEEMQGKDHFELKAWVCVSDEFDIFNISKIILQSIGGGNQEFTDLNLLQVALKEKISKKRFLLVLD
DVWSESYTDWEILERPFLAGAPGSKIIITTRKLSLLNKLGYNQPYNLSVLSHENALSLFCQHALGEDNFN
SHPTLKPHGEGIVEKCD

SEQ ID NO:18

RLG1 I

LARLVYDEMQEKDHFELKAWVCVSDEFDIFNISKIIFQSIGGGNQEFKDLNLLQVAVKEKILKKRFLLVLD
DVWSESYADWEI?ERPFLAGAAGSKIIMTTRKQSLLTKLGYKQPYNLSVLSHDSALSLFCQHALGEGNF
DSHPTLKPHGEGIVEKCAGLPLALST

SEQ ID NO:19

RLG 1J

EFGVGKTTLARLLYEEMQGKDHFELKAWVCVSDEFDIFNISKIILQSIGGGNQEFTDLNLLRVALKEKISK
KRFLLVLDDVWSESYTDWEI?ERPFLAGAPGSKIIITTRKLSLLNKLGYNQPYNLSVLSHENALSLFCQH
ALGEDNFNSHPTLKPHG?GIVEKCDGLPLALS

SEQ ID NO:20

SEQ ID NO:21

RLG2A

1 TTNACACCAT AAATTCTCNA CCTGNGGGGA CAAAAACCTA AAAATGGTCC ATAATGCNCA AATCAGNAAG
71 GTTGANAAAG CTCTAAGTTT TTNACCTCCA NCTGATGCNC NNTCCTCNTA AAGTTCANAT CCAAGCTTGC
141 CCTCCAACTC TANCNCCTTC AATGGCACCT CCTTCTCTTC AAAAGCACAC AAGAACACTT TCAAGCTCAA
211 CCACACTCAC ACAAGCTCTA GAACNAGGGT TAGGGCACAT TTAGGGTTTT GCTCTCTGGA AATGGTGTCT
281 AAAAGTGAGG CCATAATGTT CCTTATATAA GGCTCACTCC CACAATTAGG CTTTCAATCT GAACGTANTA
351 CGCCCAGTGT ACACTATGGT ACGCCCAACG TACTCGGTAG TCTCCGCGTC AANAATACAC TCATGAGTAC
421 GCGCAACGTA CTTTCCCTTA CGCCCAGCGT ACTCAAAAGC CAAACATTCT TTTCAAGGAC TAATTTTGAC
491 AACTTGAGGA AAGAAAAGGA TCAAAGANAT ATACTTGAAT TCCGGGATGT TACAATGAAG TTGANACCTT
561 GGCTAAAAAA TTAAATTGGT TGTGGAAGCC GTTGGCTGAG CAAGCAACAA GGGTAAAATT CGTAATCTAC
631 AAATGGTGTT ATTTTCTATT TCTTCTTATT ATTTTACTTG ATTTACGGGT AGTTTTTTTT TCTTACAAAA
701 AATATTAAAG TTGATAAAGT ATAGCCACTA AAATTGACTT TTTCCAAAAC ATAATGTCAA ATGGTGCGTA
771 TATGTATCAT GTTGTATTAN ATAATGAATA TGATGATNCT GTTCTATTTA ANCCGAAAAA ATTATCTAAT
841 GATTTTATAT TGGAAAACAA AGTTGTGATT TTTNGCATAA TATAATCAAA TCCNCTTTTG TNTGGGAGGT
911 GGATAAATGT GGTAAATTTA NAACAAGTGT TTTNACNTTG AAGGGTNTGG AAAGGTTGAA AAAAGTTAAA
981 ATGATAAAAT GTTTACACAA ATGTTGTATC CGACTGAATA TNATGTTTAA GGATNATTGT ATTAAATTGT
1051 TGATATATAG TAAGCATAAA TATTTAGAAT TGTGACTTAA ATTTATAAGT TATNCNAACT GGATTGAAAC
1121 ATTTTTGATA TANATTAGGA ATGAAAATGA GCAACCCTAA CATACTTATC TTTGGTAGTT TGGTTATTAT
1191 ATTTTTATTA NAATATAGAA NCATCCCTTT ATTTTAAACC CATATTGTGG ACGGACTTGA ATAAATGGGA
1261 AAAATGTACC TTGCTATTTA GCACAAAAAA ATTATAAAAA TGTACATTGC TATTTAGCAC AAACAAAAAA
1331 AAAAAACTTA TCCTTTTTGC ATTAGGTCAC AAAGAAATAT AAAATGGGAA ATGTGTTGCT ATTTAATGCA
1401 CTAAAAGAAA CTATTTTGCC TTTATTAAAC CGGGTAAACC AATAGAAAAA TGGAAGTACA TTGTCATTTA
1471 GCATGAAAAA AAATAACTTT CCATTTTTTG CATCCGGTCA CAATAATAGA AAAATGAAAG TACGTTGCTA
1541 TTTAGCGAAA CTAACTTCCT TTTTTCTTTT TGGCATCGTA TCATAAAATA TAGACTAAAA TACGTTAGTT
1611 TTACATTTTT AATACATTGA AATGTCTAAT CCACATGTTA TTCTATAAAA AGGGAAATGT AATTTACTTA
1681 TTCTTTGATT CTTTGGCTTC TTTTTAGTAC CCAAAACATC CCTCTATCCA TCTATTCCAA CTAAAATAAT
1751 GAAAACTATA TTCCTTCCAT TGTAGGGATG TTATAAATTT TGTAATTGTT TTTATGCAAA AAAGTGTTTT
1821 TTGTTAACTA GATTAACGAG ATTCATTTTT CAGCATTTTA GGAGAAGTTC ATCCATCTTT TGGATATGAA
1891 GTGCAAGCCA AGTTCTTTAA CATGGAATAT GAGGTCCCTA TATGCTCAAA AAATAGCAAA TGAGAAATTT
1961 TTTAAATTGG ATCCCCATAA AAGAAAATTT GTTAATGGTT GTTTTAATAT TGGTCAATGT GTCCACCGGA
2031 TGAGCATAAT ACTAGTTTAT AAGGGGTAAA GGTGGGTTTG GTGGGCCCAT TTATCTTTAT TATTTCTAAA
2101 AGTCAGAATT AAGTAAAAAA AATTATAAGA TAAATACCAT AAGGATAAAA AATCATTTTA TTTGGACCAA
2171 AGACCAAAGT TGTTAAGGGG CTGTTTGTTT TTTTTGTGAA GAGCTGTGCA ACCACTTTTG TCTGCGCCGC
2241 ACAGACAACG TGCAGACATA TGCCCTCGCA GAGTGTTTGT TTTTTGAAAG TGCGCAGACC AAAAAAACGT
2311 CTGCGCGAGG TCATCCTGGC GCATATATGT GTCACTGTCT TCAAAGGTCT TCAGACCTCA TTTTAACCAA
2381 AAAAAAAAAA GACCACCGGT TTTTTTTTTT TTTTTNTTCT TTCTCTTGTA GCTGAAAATG CATTTTTAAT
2451 CTTTATGACA TGAAATTAAG TTTGAAAAAT TAATTTATTT CAACAGCTGT AGACGTTAAA AACAAACAGT
2521 CTTCTTGTTG CAGACTGTGG ACATTTGGTC CACCTCTTCT ACCGCAGAGA CTTGCAGATG TGGTCCGCAG
2591 ACTGCAGACA TTTTGGCTTC AAATAAACAA ACATCACCTA ATTTGACTAC ACCACACGGA CCTCCAATGT
2661 AACAAAAAAA AGGTTGAAAC AAAGTTGCCT ATTTCTCCAT ATCCAGGGGC CATTTATGTA AGAGTTATCT
2731 AAATTTTAGT TCGGTAGATC AGTTCTCACA TTTTAACCGG GTAAAGTGTA TGTGTGTACG CGCGCACCTG
2801 AAAGGTTTGA ANGTAACTTC CAAACTGAAN CAANAATCGA TATGAAGTAT CAAGTTAGAG GTTCAATTGG
2871 TGAAGGAATC AGCTGGAGGT TGGGGAATCG AGCTTCCACT ATTAAGGTAA AATCCATAAC CCTAAATGTT
2941 GGTACGCTCA TATATCAAAT TGCGTGTTTT GTTGAATGAA AAAAGCATGC TCAAAAAACC AGTGTAAGGC
3011 ACGGTATATG ACATATTTAT AGTTACTGAT AACAAATTAT GATAATTTTG GGTTTACGTA AGTTAGGATT
3081 CGTACTTCAA CCAAATGTAA TAGTTTTTGT GAGTCTATCT ATGTATTTGG GGAATCACAT TAGCAACGGG
3151 ATTGTACTAG TAATTCGAAA AAGTCTTTTA AATAATTTTT CTGTTTATAA TTTATGAATA GTTTTAGCGA
3221 CATCTAATAT TAAATAGAAT GTATCTGATA TTGAATTAAT GTCCTTAATG TGAACATAGA CCTTTTCCAT
3291 TTACTAATGC CTAATTATTA GTTTCTAATC AATAAATTTT AATTTCTGTT TTATGCTTCT AAGACAATAA
3361 AAATCCATGA TTTACCTTTA AATATTAACA AAAATGACCA TAAATAAATA AAAAATTAGG ATACCAAACC
3431 CCCCCGCCAT GCCCAATGTC TAAATATTCT TGATGCTTTT GCTTTTCCCT CTTTTCCTTG TTAGTCTATT
3501 ATTCTGGAGA GTTTGAGAGA GTTTCATACA AGAAAATTTC AAGAAGAAAG CAAAGGTCCA GGTATTCTCT
3571 TTTCTTAATT ATGTATTAAC TTACAAGCAT TTTTTACACG ATCCATGGTT TTTTGTGTAT GTTTTTCAAA
3641 TTGAAACTAG ATTGGGACTT TTGCCCTTGA TGATTCATAA GATATTGCAT GGAGTTGAGA TTGTGTAAGA
3711 AAAGTGGTGA ATAGAAAGAG CAAGTGAATC CAGATATAGT ATTGGTAATA TATGATGATG AGATAGAGAT
3781 ATGTTAAAAC TGGCTAGAAA ATTGTTTTAA TTTGAAATTT AGGTTGTTGA ATTTGAAAGA TACCAAGCTA
3851 ATAACTAATT AGTTATGCTA AATAGTTATA AAGAACAACA AACTCGTAGT TTTTTTTTCA TGATTTTCAA
3921 CCTCTTCGTA CCAAACTAAA TTATAACAAA ATTGAATATC ATTCTCTGCA ATCAATTTTA ACTTTTGTTA
3991 TTATCATCAT GTCTAAAATT GCCACAAGTT TATTTTCATA GTCATATTGG ATTATGAAAG GACTATTTTT
4061 ACCAATTACA TCTTTACTTT ATGGCCAAAG CTAATACAAT CCGACTAAAC TAAAGGATTC TAGGATGCAT

SEQ ID NO:21

RLG 2A cont.

4131 ATAGTTTGCT CCCCGATTAT AGATTTCTAT CTAATTTGTC TATTGTACTA ATTTAGGTGC CACCACAAGT
4201 AAATTCCTGA AATGGATGTC GTTAATGCCA TTCTTAAACC AGTTGTCGAG ACTCTCATGG TACCCGTTAA
4271 GAAACACATA GGGTACCTCA TTTCCTGCAG GCAATATATG AGGGAAATGG GTATCAAAAT GAGGGGATTG
4341 AATGCTACAA GACTTGGTGT CGAAGAGCAC GTGAACCGGA ACATAAGCAA CCAGCTTGAG GTTCCAGCCC
4411 AAGTCAGGGG TTGGTTTGAA GAAGTAGGAA AGATCAATGC AAAAGTGGAA AATTTCCCTA GCGATGTTGG
4481 CAGTTGTTTC AATCTTAAGG TTAGACACGG GGTCGGAAAG AGAGCCTCCA AGATAATTGA GGACATCGAC
4551 AGTGTCATGA GAGAACACTC TATCATCATT TGGAATGATC ATTCCATTCC TTTAGGAAGA ATTGATTCCA
4621 CGAAAGCATC CACCTCAATA CCATCAACCG ATCATCATGA TGAGTTCCAG TCAAGAGAGC AAACTTTCAC
4691 AGAAGCACTA AACGCACTCG ATCCTAACCA CAAATCCCAC ATGATAGCCT TATGGGGAAT GGGCGGAGTG
4761 GGGAAGACGA CAATGATGCA TCGGCTCAAA AAGGTTGTGA AAGAAAAGAA AATGTTTAAT TTTATAATTG
4831 AGGCGGTTGT AGGGGAAAAA ACAGACCCCA TTGCTATTCA ATCAGCTGTA GCAGATTACC TAGGTATAGA
4901 GCTCAATGAA AAAACTAAAC CAGCAAGAAC TGAGAAGCTT CGGAAATGGT TTGTGGACAA TTCTGGTGGT
4971 AAGAAGATCC TAGTCATACT CGACGATGTA TGGCAGTTTG TGGATCTGAA TGATATTGGT TTAAGTCCTT
5041 TACCAAATCA AGGTGTCGAC TTCAAGGTGT TGTTGACATC ACGAGACAAA GATGTTTGCA CTGAGATGGG
5111 AGCTGAAGTT AATTCAACTT TTAATGTGAA AATGTTAATA GAAACAGAAG CACAAAGTTT ATTCCACCAA
5181 TTTATAGAAA TTTCGGATGA TGTTGATCCT GAGCTCCATA ATATAGGAGT GAATATTGTA AGGAAGTGTG
5251 GGGGTCTACC CATTGCCATA AAAACCATGG CGTGTACTCT TAGAGGAAAA AGCAAGGATG CATGGAAGAA
5321 TGCACTTCTT CGTTTAGAGC ACTATGACAT TGAAAATATT GTTAATGGAG TTTTTAAAAT GAGTTACGAC
5391 AATCTCCAAG ATGAGGAGAC TAAATCCACC TTTTTGCTTT GTGGAATGTA TCCCGAARAC TTTGATATTC
5461 TTACCGAGGA GTTGGTGAGG TATGGATGGG GGTTGAAATT ATTTAAAAAA NTGTATACTA TAGGAGAAGC
5531 AAGAACCAGG CTCAACACAT GCATTGAGCG GCTCATTCAT ACAAATTTGT TGATGGAAGT TGATGATGTT
5601 AGGTGCATCA AGATGCATGA TCTTGTTCGT GCTTTTGTTT TGGATATGTA TTCTAAAGTC GAGCATGCTT
5671 CCATTGTCAA CCATAGTAAT ACACTAGAGT GGCATGCAGA TAATATGCAC GACTCTTGTA AAAGACTTTC
5741 ATTAACATGC AAGGGTATGT CTAAGTTTCC TACAGACCTG AAGTTTCCAA ACCTCTCCAT TTTGAAACTT
5811 ATGCATGAAG ATATATCATT GAGGTTTCCC AAAAACTTTT ATGAAGAAAT GGAGAAGCTT GAGGTTATAT
5881 CCTATGATAA AATGAAATAT CCATTGCTTC CCTCATCACC TCAATGTTCC GTCAACCTTC GCGTGTTTCA
5951 TCTACATAAA TGCTCGTTAG TGATGTTTGA CTGCTCTTGT ATTGGAAATC TGTCGAATCT AGAAGTGCTT
6021 AGCTTTGCTG ATTCTGCCAT TGACCGGTTG CCTTCCACAA TCGGAAAGTT GAAGAAGCTA AGGCTACTGG
6091 ATTTGACGAA TTGTTATGGT GTTCGTATAG ATAATGGTGT CTTAAAAAAA TTGGTCAAAC TGGAGGAGCT
6161 CTATATGACA GTGGTTGATC GAGGTCGAAA GGCGATTAGC CTCACAGATG ATAACTGCAA GGAGATGGCA
6231 GAGCGTTCAA AAGATATTTA TGCATTAGAA CTTGAGTTCT TTGAAAACGA TGCTCAACCA AAGAATATGT
6301 CATTTGAGAA GCTACAACGA TTCCAGATCT CAGTGGGGCG CTATTTATAT GGAGATTCCA TAAAGAGTAG
6371 GCACTCGTAT GAAAACACAT TGAAGTTGGT TCTTGAAAAA GGTGAATTAT TGGAAGCTCG AATGAACGAG
6441 TTGTTTAAGA AAACAGAGGT GTTATGTTTA AGTGTGGGAG ATATGAATGA TCTTGAAGAT ATTGAGGTTA
6511 AGTCATCCTC ACAACTTCTT CAATCTTCTT CGTTCAACAA TTTAAGAGTC CTTGTCGTTT CAAAGTGTGC
6581 AGAGTTGAAA CACTTCTTCA CACCTGGTGT TGCAAACACT TTAAAAAAGC TTGAGCATCT TGAAGTTTAC
6651 AAATGTGATA ATATGGAAGA ACTCATACGT AGCAGGGGTA GTGAAGAAGA GACGATTACA TTCCCCAAGC
6721 TGAAGTTTTT ATCTTTGTGT GGGCTACCAA AGCTATCGGG TTTGTGCGAT AATGTCAAAA TAATTGAGCT
6791 ACCACAACTC ATGGAGTTGG AACTTGACGA CATTCCAGGT TTCACAAGCA TATATCCCAT GAAAAAGTTT
6861 GAAACATTTA GTTTGTTGAA GGAAGAGGTA AATATAAATT TTTAATGCTA ATACATTACA AAGGATCTTT
6931 TCAGTTAAAT CTTTCAAAAT ATATTGTAAT TTGATTGTAT GGGGTATTAT TGTTGGATGG GACTATTAAT
7001 AAATGATTAT CTTGCAGGTT CTGATTCCTA AGTTAGAGAA ACTGCATGTT AGTAGTATGT GGAATCTGAA
7071 GGAGATATGG CCTTGCGAAT TTAATATGAG TGAGGAAGTT AAGTTCAGAG AGATTAAAGT GAGTAACTGT
7141 GATAAGCTTG TGAATTTGTT TCCGCACAAG CCCATATCTC TGCTGCATCA TCTTGAAGAG CTTAAAGTCA
7211 AGAATTGTGG TTCCATTGAA TCGTTATTCA ACATCCATTT GGATTGTGTT GGTGCAACTG GAGATGAATA
7281 CAACAACAGT GGTGTAAGAA TTATTAAAGT GATCAGTTGT GATAAGCTTG TGAATCTCTT TCCACACAAT
7351 CCCATGTCTA TACTGCATCA TCTTGAAGAG CTTGAAGTCG AGAATTGTGG TTCCATTGAA TCGTTATTCA
7421 ACATTGACTT GGATTGTGCT GGTGCAATTG GGCAAGAAGA CAACAGCATC AGCTTAAGAA ACATCAAAGT
7491 GGAGAATTTA GGGAAGCTAA GANAGGTGTG GAGGATAAAA GGTGGAGATA ACTCTCGTCC CCTTGTTCAT
7561 GGCTTTCAAT CTGTTGAAAG CATAAGGGTT ACNAAATGTN AGAAGTTTAG AAATGTATTC ACACCTACCA
7631 CCACAAATTT TAATCTGGGG GCACTTTTGG AGATTTCAAT AGATGACTGC GGAGAAAACA GGGGAAATGA
7701 CGAATCGGAA GAGAGTAGCC ATGAGCAAGA GCAGGTAAGG ATTTCAATTT CACTGTCTTA ATTAATGATT
7771 AAGCTCCTGC TTTTTGAATA AAAAAGGGAC AAACCATTTC ATGACTTAAT GTAGCAATAC AAGTCATGTA
7841 TAAGAGTGAC CAACTCTTTT TTATTTATAA AATGACTACA AAATATTTTT TTTCATTAGA GATCATGTAT
7911 AAATGTGACT AATTTTTCAT CACCTAACTT TAGTTGATAA ATCTTTATAA ATGTCACTAG TTACTTTTCA
7981 GTAAAATAAC AAATTTAATA AATTATCAAC AAAAAGCATC AACTAAAAAA ATCCCACAAC CCGTAATAAT
8051 TTAAAATAAA AGGATTTAAC ATCTAATACG AACAATTTTT TTTCTAAACA TGATTTGGAC CAAATATCAC
8121 CAGCAACTCA AGTTTGGAAT CGATTCAGCT TAAAACTTGA CCAGCATAAT TAGATAGATG AGAGTTGAAG
8191 CTAAAGTGCC TATATAAGTT CGTTTCATCT TTTTTCTTGA TCTTGATAGC AAGTTGAATG ATTTTCTTCT

RLG2A cont.

8261 TCAAAATTGA TAAAAATCTA CATTATAAAG AGACTAGCTT GAAAAAAAAT GGTCTAGGTG GGTCTTGGGT
8331 TCTGGTAGAT GAAGATGGAA GGGGAGAGTA TGATTTCAAA GACACAACAC ATCCTTCATT TTATTTATTT
8401 ATTATTATTA TTATTTTTTG ATATCTTGCT CATATTTGTT ACAGATATGT GAGGTCTATT AATCTTTTTA
8471 AATATATAAA AAAATAAATA ACATAAATGA GAAAATTAAA TAAAGAATAA ATTAATAAGG GCACAATAGT
8541 CTTTTTAGGT AAGACAAGGA CCAAACACGC AACAAAAATA AACAGTAGGG ACCATCCGAT TTAAAAAAAA
8611 TAATTAGGGA CCAAAAACAT AAATTCCCCC AAACCATAGG GACCATTCAT GTAATTTACT. CTTACTTTTC
8681 GTTTTGTTCA TATTTGGGTA ACTATTTTTT TTGTACACAT CTAGGTAACG AACTTGTTGA AGTGTTCCCA
8751 TTTAGGATGT GACCTACTAC AACCGATCAT AATAGTCATA TGTGAACACT TCCAACAACT TTATTACTTA
8821 GGTGTGTACA AAAAAACAAT AGTTACCATG ATGTGAACAT ACTGAAAAAT TAATTACCTT AGCAAGTTAT
8891 TTTCCCATTT AGGTTGTATG GAAACAGTTC CGTGAGACCG TGACTTGGAT GGTAGATAAA TTTAGTAAAC
8961 TTAACCCTTC AATTAACCTA CCTTTTTCTT ATTAACTCAA TTTCAACCTA AATTCTGATT CTTGTTTGAA
9031 AGTAAGTTGC ATCTTTATTT TTGTATTATC TTGTTGCATA GGATCCTTAG CATCTTTTAA TAATTTATTT
9101 GAAGGTGAAA GATCCAACTA TTTTTAATCT GTTGGCATTT TCCATCATTT GCAACTGTTT CTTGAAAAAA
9171 AAATACCTAA AATCAAAATA ACCATTTTCA AATCCAAAAT TATAAGAGAG AATTGTAAAT GGACATGGAA
9241 TCATAAATCA TTAACACAGT TCAGTAAACA AGTTGCTAAT TACATTTCTT GCTGTGCAGA TTGAAATTCT
9311 ATCAGAGAAA GAGACATTAC AAGAAGCCAC TGACAGTATT TCTAATGTTG TATTCCCATC CTGTCTCATG
9381 CACTCTTTTC ATAACCTCCA GAAACTTATA TTGAACAGAG TTAAAGGAGT GGAGGTGGTG TTTGAGATAG
9451 AGAGTGAGAG TCCAACAAGT AGAGAATTGG TAACAACTCA CCATAACCAA CAACAACCTA TTATACTTCC
9521 CAACCTCCAG GAATTGATTC TATGGAATAT GGACAACATG AGTCATGTGT GGAAGTGCAG CAACTGGAAT
9591 AAATTCTTCA CTCTTCCAAA ACAACAATCA GAATCCCCAT TCCACAACCT CACAACCATA AAAATTATGT
9661 ATTGCAAAAG CATTAAGTAC TTGTTTTCGC CTCTCATGGC AGAACTTCTT TCCAACCTAA AGCATATCAA
9731 GATAAGAGAG TGTGATGGTA TTGGAGAAGT TGTTTCAAAC AGAGATGATG AGGATGAAGA AATGACTACA
9801 TTTACATCTA CCCACACAAC CACCACTTTG TTCCCTAGTC TTGATTCTCT CACTCTAAGT TTCCTGGAGA
9871 ATCTGAAGTG TATTGGTGGA GGTGGTGCCA AGGATGAAGG GAGCAATGAA ATATCTTTCA ATAATACCAC
9941 TGCAACTACT GCTGTTCTTG ATCAATTTGA GGTATGCTTT GTACATATTC AATTATTTAT TTAATTTCCT
10011 TTTTTATTTG CAATATTCTA TAAATAATAC ATTTTATACC CACTATACTA AGATAATAAT TACCTAGAGG
10081 GATGGATGCT ATGACACAGC TGCTACACTT CAGAAACTCT AGTAAGGGCA GTTATGGAAG TTCAATAAAA
10151 TGATAATGGC ATCTTTTGAT GGGTAATATA GGCAATTTAA GTTTTATTTC TGTTAAAGCA GTATTTAGCA
10221 AGTACTGGCC AGTAGGAGAG GAGAATATCA CCTTTTGTGA AAATCTGGTC ATTGTACCCA GAATTTAGTT
10291 AAATGTAACA TTTTAGATAT CAGGGGTCAT CAGGTGACAG ATATTGTAGA ATAGAACAAT ATATAATATC
10361 ACCCAAAACT ATTTTTTCTA AGGTTATTCT GTTAAATATG TGCTTTCTTG TTTTCATNGA ATTNGCATTC
10431 GTATATTTTA GGTGTTAAAG TGATTTTNTC TTCAATAAAT CCCGAAATTA ATTAAAAAAA AAAAAACAAA
10501 AGTACATTTT TGATGTGGAG AGCACTGGTA TCACTTAGTA TATAAAAAGC TTGATTTTGA ATTAACTTTC
10571 TTATACAAAA GTTGTGTATA TAGTTTAATT AGTTTTACAT CATTTTTCCA TGTGGTGTTG CAGTTGTCTG
10641 AAGCAGGTGG TGTTTCTTGG AGCTTATGCC AATACGCTAG AGAGATGAGA ATAGAATTCT GCAATGCATT
10711 GTCAAGTGTA ATTCCATGTT ATGCAGCAGG ACAAATGCAA AAGCTGAAGG AGAGGACAGC GATTCTCGTA
10781 CGAACGGTTA CGATTCGACT GGCCGTCGTT TTACA

SEQ ID NO: 21

RLG2A a.a.

MDVVNAILKPVVETLMVPVKKHIGYLISCRQYMREMGIKMRGLNATRLGVEEHVNRNISNQLEVPAQV
RGWFEEVGKINAKVENFPSDVGSCFNLKVRHGVGKRASKIIEDIDSVMREHSIIIWNDHSIPLGRIDSTK
ASTSIPSTDHHDEFQSREQTFTEALNALDPNHKSHMIALWGMGGVGKTTMMHRLKKVVKEKKMFNFII
EAVVGEKTDPIAIQSAVADYLGIELNEKTKPARTEKLRKWFVDNSGGKKILVILDDVWQFVDLNDIGLS
PLPNQGVDFKVLLTSRDKDVCTEMGAEVNSTFNVKMLIETEAQSLFHQFIEISDDVDPELHNIGVNIVRK
CGGLPIAIKTMACTLRGKSKDAWKNALLRLEHYDIENIVNGVFKMSYDNLQDEETKSTFLLCGMYPE?FD
ILTEELVRYGWGLKLFKK?YTIGEARTRLNTCIERLIHTNLLMEVDDVRCIKMHDLVRAFVLDMYSKVEH
ASIVNHSNTLEWHADNMHDSCKRLSLTCKGMSKFPTDLKFPNLSILKLMHEDISLRFPKNFYEEMEKLE
VISYDKMKYPLLPSSPQCSVNLRVFHLHKCSLVMFDCSCIGNLSNLEVLSFADSAIDRLPSTIGKLKKLR
LLDLTNCYGVRIDNGVLKKLVKLEELYMTVVDRGRKAISLTDDNCKEMAERSKDIYALELEFFENDAQPK
NMSFEKLQRFQISVGRYLYGDSIKSRHSYENTLKLVLEKGELLEARMNELFKKTEVLCLSVGDMNDLEDIE
VKSSSQLLQSSSFNNLRVLVVSKCAELKHFFTPGVANTLKKLEHLEVYKCDNMEELIRSRGSEEETITFP
KLKFLSLCGLPKLSGLCDNVKIIELPQLMELELDDIPGFTSIYPMKKFETFSLLKEEVLIPKLEKLHVSSM
WNLKEIWPCEFNMSEEVKFREIKVSNCDKLVNLFPHKPISLLHHLEELKVKNCGSIESLFNIHLDCVGAT
GDEYNNSGVRIIKVISCDKLVNLFPHNPMSILHHLEELEVENCGSIESLFNIDLDCAGAIGQEDNSISLRNI
KVENLGKLR?VWRIKGGDNSRPLVHGFQSVESIRVTKC?KFRNVFTPTTTNFNLGALLEISIDDCGENR
GNDESEESSHEQEQIEILSEKETLQEATDSISNVVFPSCLMHSFHNLQKLILNRVKGVEVVFEIESESPTS
RELVTTHHNQQQPIILPNLQELILWNMDNMSHVWKCSNWNKFFTLPKQQSESPFHNLTTIKIMYCKSIKY
LFSPLMAELLSNLKHIKIRECDGIGEVVSNRDDEDEEMTTFTSTHTTTTLFPSLDSLTLSFLENLKCIGGG
GAKDEGSNEISFNNTTATTAVLDQFEVCFVHIQLFI.

SEQ ID NO:22

RLG2B

SEQ ID NO:23

1 AGTTTTTTTT TTTCCCAATA TCCATTTATA TGCGATTTAT TTCTGAAATA ATTTTATCAA AACGCAGGAA
71 ACAATGTAGA ATAATACTGG TATAATTAAT TATATAAAGT TATTAGGCTG AAATCTTGAG GCTACTATAA
141 TTTAATTATC ATAATTTGAA AATCATCAAA TTGTATTCCA TGTATATTTA TGTTATCAGA TAATTAATAA
211 TATGTGAGCC ACACAAATCC ACATCATCAG ACACCCCACC TTATTGTCGG CTACCTCACC ACTTGCATGA
281 TCCCGACATC TTCCCAACCC CACCGACGAC TTGGGGTCTC CTTAATATAT CAATTATTTT CTGTAAGTAT
351 TTATTTGTGT AAATGTGTAA TGTCATTTTA CCTTTTTTCT AATATATACA GAAACATAAA TTTTAAATGA
421 AATTCAACTG CGTTTCATTC TTGCATTAAA AAAAAAGACT GTACTGTTGT CAATATTTTA CTTATAACCT
491 GATTAATTAA TTAAAGCGTA ATTGCATAAT TTGCATTAGG TTGTAATTTT GTGTTTTATA GGGAGGGTGA
561 GGGTCACCGG GAATCAAAGC ACTTATGTAA AAGCAGGGGA AATACAAAAA ATTTACTCGA AACAAATTTT
631 ATTCAATTTA AGTGAGATAA TAATGTTCTG ATTAGATTAT GAGAACTAGG AGATTTAAGT GATATATCCC
701 ATTTAAAAGA AATTGCATTA TTAATTTTGG ATCTCTTGAT GATGACAAAA TTAACTCGTG ACAGGTTATA
771 TATCATATAC AAAATGAGTG GCTATGCTTT CGCTTTCCAA AAAGCAATTA TAGTTATACT ACACCTACAA
841 ATTTTAAAAG GGGTTAAACA TATCAAAATA CTTGATAAGT AATTATATAA ATATGCATTT AACCCTCTAA
911 AGAAAATGCT ACTAAGCTTG GACCATCTCA GAATTACAAT CATACCCTTC CCCTCAAAAA AGATTCGTAT
981 ATATCATGTC ATTTGGCATT CATTTCTTTT TCACAATTCA TAGTTCTATT CTCAAAAAAT TCGAGTTCTC
1051 GTATTTGTAA GGAAGATCAG AAGAGACTGT TCACACAGGT ACTCTCTTTT ATTTATTGAT TCACATTCAT
1121 ATATGTTATT GTTTTCTTGC TTAATGGTTT CGTCAGTCTA ACTGCGCTTG CTGATTTAAA TTTCTTCACT
1191 TTCTTCCACG GATTTTTTAA ATATTAGTTT TGTGAATGAA CAATTGGTGA AGGAAAGAAA CATGGGAGTC
1261 TTTTCTAAAG TAAACCTAGA TACTTAGGTT ATAAGGGTAT ATGCTAAAAT GAACTATGCC CATTCACCTT
1331 TGCCTTTTCT TTTACTTTTT AGTTTTTAGA ATCCAAGTTT TCATATGTAT CTCGATGTGT GAGAAGAATA
1401 GGCATTAGAA AGGTAAAGGA CGTACATAAA ATTGATTAAT TAGTGAATGT TCTTTGATAT CATTATTTTT
1471 ACTCTCATAA AAAGCATATA GATCAAACAC AAATTGCTAC TTGTTAGTGT AACAACTTCG ACTTAATAAT
1541 GTTAATAATC AAGATTCTCT TGATTTCAAC TATTTTCTAA CCGAACAAGC TCACTAAAAA CTCATATTGC
1611 TTTGAGTCTG AGTGGTTTAT ATTTGGGGTT TTACATTTAA TTTTTTGTGC ATGAATGTGA AAATAGACTG
1681 CTTATTGATT CTTTGTGTTT CATTGAGTTG ATTTTCATTA TTACTACCTT ACAAATTGCT CAGTGATAGA
1751 TTTCCATTAA TTTGCTAATT CGGTTGCTTC TAAATATGTA GGAGCTACTA AAAGCAAAAA TATCGAGCAA
1821 TGTCGGACCC AACGGGGATT GCTGGTGCCA TTATTAACCC AATTGCTCAG ACGGCCTTGG TTCCCGTTAC
1891 GGACCATGTA GGCTACATGA TTTCCTGCAG AAAATATGTG AGGGTCATGC AGATGAAAAT GACAGAGTTG
1961 AATACCTCAA GAATCAGTGT AGAGGAACAC ATTAGCCGGA ACACAAGAAA TCATCTTCAG TTCCATCTCA
2031 AACTAAGGAA TGGTTGGACC AAGTAGAAGG GATCAGAGCA AATGTGGAAA ACTTTCCGAT TGATGTCATC
2101 ACTTGTTGTA GTCTCAGGAT CAGGCACAAG CTTGGACAGA AAGCNTTCAA GATAACTGAG CAGATTGAAA
2171 GTCTAACGAG ACAACTCTCC CTGATCAGTT GGACTGATGA TCCAGTTCYT CTAGGAAGAG TTGGTTCCAT
2241 GAATGCATCC ACCTCTGCAT CATTAAGTGA TGATTTCCCA TCAAGAGAGA AAACTTTTAC ACAAGCACTA
2311 ATAGCACTCG AACCCAACCA AAAATTCCAC ATGGTAGCCT TGTGTGGGAT GGGTGGAGTG GGGAAGACTA
2381 GAATGATGCA AAGGCTGAAG AAGGCTGMTG AAGAAAAGAA ATTGTTTAAT TATATTGTTG GGGCAGTTAT
2451 AKGGGAAAAG ACGGACCCCT TTGCCATTCA AGAAGCTATA GCAGATTACC TCGGTATACA ACTCAATGAA
2521 AAAACTAAGC CAGCAAGAGC TGATAAGCTT CGTGAATGGT TCAAAAAGAA TTCAGATGGA GGTAAGACTA
2591 AGTTCCTCAT AGTACTTGAC GATGTTTGGC AATTAGTTGA TCTTGAAGAT ATTGGGTTAA GTCCTTTTCC
2661 AAATCAAGGT GTCGACTTCA AGGTCTTGTT GACATCACGA GACTCACAAG TTTGCACTAT GATGGGGGTT
2731 GAAGCTAATT CAATTATTAA CGTGGGCCTT CTAACTGAAG CAGAAGCTCA AAGTCTGTTC CAACAATTTG
2801 TAGAAACTTC TGAGCCCGAG CTCCAGAAGA TAGGAGAGGA TATCGTAAGG AAGTGTTGCG GTCTACCTAT
2871 TGCCATAAAA ACCATGGCAT GTWCTCTTAG AAATAAAAGA AAGGATGCAT GGAAGGATGC ACTTTCGCGC
2941 ATAGAGCACT ATGACATTCA CAATGTTGCG CCCAAAGTCT TTGAAACGAG CTACCACAAT CTCCAAGAAG
3011 AGGAGACTAA ATCCACTTTT TTAATGTGTG GTTTGTTTCC CGAAGACTTC GATATTCCTA CTGAGGAGTT
3081 GATGAGGTAT GGATGGGGCT TGAAGCTATT TGATAGAGTT TATACGATTA GAGAAGCAAG AACCAGGCTC
3151 AACACCTGCA TTGAGCGACT GGTGCAGACA AATTTGTTAA TTGAAAGTGA TGATGTTGGG TGTGTCAAGA
3221 TGCATGATCT GGTCCGTGCT TTTGTTTTGG GTATGTTTTC TGAAGTCGAG CATGCTTCTA TTGTCAACCA
3291 TGGTAATATG CCTGGGTGGC CTGATGAAAA TGATATGATC GTGCACTCTT GCAAAAGAAT TTCATTAACA
3361 TGCAAGGGTA TGATTGAGAT TCCAGTAGAC CTCAAGTTTC CTAAACTAAC GATTTTGAAA CTTATGCATG
3431 GAGATAAGTC GCTAAGGTTT CCTCAAGACT TTTATGAAGG AATGGAAAAG CTCCATGTTA TATCATACGA
3501 TAAAATGAAG TACCCATTGC TTCCTTTGGC ACCTCGATGC TCCACCAACA TTCGGGTGCT TCATCTCACT
3571 GAATGTTCAT TAAAGATGTT TGATTGCTCT TCTATCGGAA ATCTATCGAA TCTGGAAGTG CTGAGCTTTG
3641 CAAATTCTCA CATTGAATGG TTACCTTCCA CAGTCAGAAA TTTAAAGAAG CTAAGGTTAC TTGATCTGAG
3711 ATTTTGTGAT GGTCTCCGTA TAGAACAGGG TGTCTTGAAA AGTTTTGTCA AACTTGAAGA ATTTTATATT
3781 GGAGATGCAT CTGGGTTTAT AGATGATAAC TGCAATGAGA TGGCAGAGCG TTCTTACAAC CTTTCTGCAT
3851 TAGAATTCGC GTTCTTTAAT AACAAGGCTG AAGTGAAAAA TATGTCATTT GAGAATCTTG AACGATTCAA
3921 GATCTCAGTG GGATGCTCTT TTGATGAAAA TATCAATATG AGTAGCCACT CATACGAAAA CATGTTGCAA
3991 TTGGTGACCA ACAAAGGTGA TGTATTAGAC TCTAAACTTA ATGGGTTATT TTTGAAAACA GAGGTGCTTT
4061 TTTTAAGTGT GCATGGCATG AATGATCTTG AAGATGTTGA GGTGAAGTCG ACACATCCTA CTCAGTCCTC

RLG2B cont.

SEQ ID NO:23

4131 TTCATTCTGC AATTTAAAAG TTCTTATTAT TTCAAAGTGT GTAGAGTTGA GATACCTTTT CAAACTCAAT
4201 CTTGCAAACA CTTTGTCAAG ACTTGAGCAT CTAGAAGTTT GTGAATGTGA GAATATGGAA GAACTCATAC
4271 ATACTGGAAT TGGGGGTTGT GGAGAAGAGA CAATTACTTT CCCTAAGCTG AAGTTTTTAT CTTTGAGTCA
4341 ACTACCGAAG TTATCAAGTT TGTGCCATAA TGTCAACATA ATTGGGCTAC CACATCTCGT AGACTTGATA
4411 CTTAAGGGCA TTCCAGGTTT CACAGTCATT TATCCGCAGA ACAAGTTGCG AACATCTAGT TTGTTGAAGG
4481 AAGGGGTAGA TATATGTTCT TTATGTTAAT ACAATTTAAA TAATATTTTC AACCAAATTT TCATAATATA
4551 TCTGTAATTT GATTGTATGA TGTGTTATTG TTTATATGTG GCTATTAAGG GATGATTATT TTGCAGGTTG
4621 TGATTCCTAA GTTGGAGACA CTTCAAATTG ATGACATGGA GAACTTAGAA GAAATATGGC CTTGTGAACT
4691 TAGTGGAGGT GAGAAAGTTA AGTTGAGAGC GATTAAAGTG AGTAGCTGTG ATAAGCTTGT GAATCTATTT
4761 CCGCGCAATC CCATGTCTCT GTTGCATCAT CTTGAAGAGC TTACAGTCGA GAATTGCGGT TCCATTGAGT
4831 CGTTATTCAA CATTGACTTG GATTGTGTCG GTGCAATTGG AGAAGAAGAC AACAAGAGCC TCTTAAGAAG
4901 CATCAACGTG GAGAATTTAG GGAAGCTAAG AGAGGTGTGG AGGATAAAAG GTGCAGATAA CTCTGATCTC
4971 ATCAACGGTT TTCAAGCTGT TGAAAGCATA AAGATTGAAA AATGTAAGAG GTTTAGAAAT ATATTCACAC
5041 CTATCACCGC CAATTTTTAT CTGGAGGCAC TTTTGGAGAT TCAGATAGAA GGTTGCGGAG GAAATCACGA
5111 ATCAGAAGAG CAGGTAACGC TTTCAATTTC ACTTTCTTAA TTAATTAAGG ACTAAGCTCC TGTTTTTTGA
5181 ATAATAAAGA GGTGGGATGA CTAAACTTGG GCATCACAAT TGCAACAAAA TGTTACAAAC CATGAAACGT
5251 TCAAACCATT TCTTGAATTA AGGTTTCAAT ACAAGTCATT TAAAAATATG GCTTAAATTT TTTTTATATT
5321 TATGTATCAA CATGATTTTT CATTAGAGAT CATTATTATA ATAGTAAGTT TAAAGCAATT TAAATCAGAA
5391 CTAATTCTAA CTTTAGCTAA TAAATCGTTA TAAATGTAAA TAATTACTTT TTAGTGAAAT AAGCAACGGA
5461 TTTAATAAGT TAACAACTTA AATGTCATTT CCTAACAAAA AAAACTTTGG TTCAGAAAAA CCGCAATTCA
5531 AGATAACTAA AATAAAAATA TTTGACATTC ACTAAGAGCA TTTTTTTTTC TAAATATGAT TGCAAATGAA
5601 TAAAACTTAA ATTTATACAG AAAATTCTTT TATATATGTT ATACAAAATT TACAAATTGA AATTGGATAT
5671 GTTAATTAAC GGTTTATAAT TCTGGTATCA CAAAGGGATA TATAATAAAA TATTATTTTC TGTAGTCATT
5741 TGTAATTGTA CTAGTTTATA ACCCGTGGGA ACCATGAGTT CTAAAATTAG TTAAACTTTC ATAATAAAAA
5811 TTTATAATTA TTATTTATTT TAAATAAATT ATTAATTAAG AGATATATCA AAAATTTAAA GTTATTATAA
5881 CTTCAAATTT AACATATAAT TAGAAAATAT ATGATCATAA CTTCTGCACT CTCTTTGTAT AAATGCAGAG
5951 AAGCTATTAG TATATTTCTA ATCAAGTCCA AACCTAATGA AGCCTATATA ATTTTGTGAA AACTCAATTA
6021 GCATTAGGTT TTAAGAGTCA CCAAATTCAA AGAATAATCC AATGCTTTCA TTACCACTAT GGAGAAAATA
6091 TTTTCTTAGT TTAAATGAAA TGAAAACAAA CATTCAAACT AATTGTTGCT TATTAAACCA AAGACCCATT
6161 ACTTAGCCAA GAGTTTAACA AAAAAAAATT ACATTCATGT ATCATTATTC ATGACTAGAT ATATATGAAC
6231 ATGAAGGGAG TTTTTATAGA AAATATAATC ATAGATATTC AACATAACTT CAGGGAATTC CTCAAAATAA
6301 CCAAGTTATT CAAGAAATTA CATCCAAGTC AACCAAAGAG AAGTTTAGCC TAGCATGGCT AAACTCAAGA
6371 AACTAAAATA AGGATTAGAA GTACCAAACA TGTAGTAAGA ATCACAGTAA AAGATGATGT TGTTCTTGAT
6441 GTTCTTCTAA GTTCTTCAAG TCTCCAGTTG CTCCTAATAA TGCAAAGGAG AGCCATTAAA TTCGTATGTA
6511 TTGATCCCTT CAAAAGCTGC ACCAACCTCC CTTAAATAAC ACTCAAAGCA AAAATGACAA AATGCCCTGA
6581 AGGACCCTAT GTGGGTGCCT TGCGCGGGTG GAGCTGCATA CGAAAGGTCT TTGGTCTTTG TGAGGGTGAT
6651 GTTGTGCGGG ATAGCTTGTC GCATGCTTCC GCGCGGTTCA CGCACATGTG CACAGGTGAT GCATGGTGTG
6721 TGCGTTCTTG AGTTTTGAGC CTCCGATGCT TAGTCCACTT GGCCCAATTC GAGTCCAATC AGCTTATAAC
6791 CCATTTTTCT TCAAGTTATC TTCAAGTTAA GCCCAATTTG GCTTCTCCAA ATCATCCATA ACTTCACAGA
6861 ATCGCCCGTT CATCTTAATC CCGGATGCAC AATTATTCTC CCGTCTTCAT TTTAAGCAAG ATACCACCTT
6931 CTTCATGCTT CATCCATCAA TAGTACACTT CATGTATCAT CTCTACTAGT TATTTAGTCC ACAAATCCTT
7001 GTTGTCCTCC AAATTTAATT ATCTCATTTA GTTCCCCGTT CCGCTACTTT CCTTAAAATT TGGAATTAAG
7071 CTCAGAGAAA TATTAAGTAC CCGAAATGGT CATAAAATTA ACAAAAAGGA AAATGCATGA AGATTAACTA
7141 AATGATGAAC GAAATATGCT AAAATAGACT ATAAAATGAA GTAAATAAAA TGAAATTATC GCACTCCGAC
7211 CACCCTTATG GCTTGTAGTC CACCCACCCT TCATTCCTTG TACCAATATG GGATGGAAAC ATCATTAATT
7281 AAGCCAAAAA GCTAACATAT AAGGGTTTAG TGACAAAGGT AAGTACTAAA GATGAAAATA ATCCATTTTT
7351 CTTGTTTTTA CACAACACAC ACATAGGGGC AGACGTAGGA TTTCAAAGTA CAGATTGTTG GTGGCACATA
7421 AGTGTTGCTG GTGACATTTT TTTTTTCTTT TTACGTGGTG GCACAACAGT AGGAAAAACG AAAAATTCGA
7491 AATTTTTTAC AATTTGTCTT AAAAAAAACA GGGGTTGTTG GTGCCACTAT GGACAACAAA GTTGAACTGC
7561 CCTACGCGCG CACACACACA CACACACATA GAGAGAGAGA GAGAGAGAGA GAGAGAGAGA AAGAAAGAAA
7631 GAGAGAGAGA GTTTGGGATG TGATACTTCT TTTAGGAAAA TGGAGTTATA TCTTTGATAT TGTATTTTTT
7701 TAATGTAATT TATNTATTTA ATCATTTTAG TTTATAAGTT NTATTTATTN GGNTATGAAA AAAAAAGTCT
7771 TTTATACATT GGATTTAACA TAAAAATCCA ACAATATTAA TCAAAAAGAC CAAACATGTG GACAATTATG
7841 TATATAATTA ATTCACAATA GTCTTTAGGA ATAGTATTAT ATATATAATT AATTCTCAAT GGTCTTAGGA
7911 ATAGTAAGTT CTTATATTTC AAACTTTTGC CACAATTCTT TGCTTACTTT GACACTTTTC CTTCCTAACT
7981 TTACATATAT ATATATATTA AAGCGCAAAG GTCATAGGAA TATAATATTT TCTATTATTC TACGTTTTGC
8051 CACAAAAGTT TGAACACTTT GCCACTTTTT GTCCCTCCTT AACCTTTTCA ATGTTTTGCG ACAAAAGTTC
8121 CAAAACTTTG CCACTTTGAT CATTCCTCAA CTTTTCACCG CATTAGTTTG TGGAGTTGGC AGTTTTGGTC
8191 CCTCTAACTT CGATATTCTC TACTGCTAGC CAAAAAGGGT TCCAGAGTTT CACACTTTTG GTCCCTGACA

RLG2B cont.

8261 GTAACCAAAT GTGAGATGTC AAATTTTTGC CACATTAGTT TGTGGAGTTG TCCCTTTTGG TCCCCCCACA
8331 TTCGATATTC TACTATACGA TCTTATTTTT CTCAAATAAC AACACGTATA TTTCATC:CT AATTGGAAAA
8401 AGAGTTTTAA AA:AAATAAC GACTAGG::: G:GC:GAGTT TTTTTT:ACA AGTTTGTATC AAATCATATC
8471 AAAATTTAAG GTGGAACGGT GACCACATTA ACCAGAAATG TAATTTATTC TTTGATTTTG ATAATTTTTA
8541 ATATTTTGTT GTGATCTATG TATTTAAAAG TAAACAACAA AGAACATAAT CCAAAACCCT AAATTGCAAG
8611 TCTCGCCCAA TTTCTCTATC ACTAGTCCTC ACTTACGATG GCGTTACGTC GCTCTCTCAC TGCTTACAAC
8681 CCTTTGTTGC TACTCATTAC AATAACGAAA AGTTGAATAT CCATATATTT ATTTGGATGT GGAATTGAAC
8751 GAATCTCGTC AAAATTTTGA TTTTGTTGAT GGATTTGAGT AGAAGTTTGG GCAGAACGGG AATGATGGTC
8821 TGCAAGTGGT TATAAACTTG ATTCTGAGTT ATTACTATAT ATGTAGCCTC TTTACAACGA CCAAGGTTTC
8891 TTCCAGGTAC CATTTGATCT TTTTAGAACT TAGTTTTCTG AAACACCCTG ATTTGGATCA AATATCACCA
8961 ACAACTCTTA AAAACTTGAT TAATCAATTG TTTTCTTCAT CTTGATAACA AGTGGAATGA TTTTCTACTT
9031 AGATTAACTT GAAAAAAAAG GTCCATGTGC GTCTGGTGGA TCTGGTAAAT GAAGATGGAA GGGAGAGCTG
9101 ACTTTAAAGA CACAAACACG TCACCATATC TCTTATTTTA TTTTAAATTT GCTTTTGGTG TATTTTCTTT
9171 TTTCCTATTT CTTTCTTTCT TGATCTCCAG ATGGTATGTG GTGTGGATAA TTTACACCTA GAGATTGGGA
9241 ACGATGGGAA GGGGTCTGTG ATTTATGGCT GGCCGAGTTT TACTTATTAA CTCAATTTCA ACCTAAATTC
9311 TGATTCTTGT TTGAAAATAA GTTGCATCTT TATTTTTGTA TTATCTTGTT GCATAGGATC CTTAGCATCT
9381 TTTAATAATT TATTTGAAGG TGAAAGATCC AACTATTTTT TAGCTGTTGG CATTTTCCAT CATTTGCAAC
9451 TGTTTCTTGA AAAAAAAATA CCTAAAATAA AAATAACCAT TTTCAAATCC AAAATTATAA GAGAGAATTG
9521 TAAATGGACA TGGAATCATA AATCATTAAC ACAGTTCAGT AAACAAGTTG CTAATTACAT TTCTTGCTGT
9591 GCAGATTGAA ATTCTATCAG AGAAAGAGAC ATTACAAGAA GCCACTGGCA GTATTTCAAA TCTTGTATTC
9661 CCATCCTGTC TCATGCACTC TTTTCATAAC CTCCGTGTGC TTACATTGGA TAATTATGAA GGAGTGGAGG
9731 TGGTATTTGA GATAGAGAGT GAGAGTCCAA CATGTAGAGA ATTGGTAACA ACTCGCAATA ACCAACAACA
9801 GCCTATTATA CTTCCCTACC TCCAGGATTT GTATCTAAGG AATATGGACA ACACGAGTCA TGTGTGGAAG
9871 TGCAGCAACT GGAATAAATT CTTCACTCTT CCAAAACAAC AATCAGAATC CCCATTCCAC AACCTCACAA
9941 CCATAAATAT TCTTAAATGC AAAAGCATTA AGTACTTGTT TTCGCCTCTC ATGGCAGAAC TTCTTTCCAA
10011 CCTAAAGGAT ATCCGGATAA GTGAGTGTGA TGGTATTAAA GAAGTTGTTT CAAACAGAGA TGATGAGGAT
10081 GAAGAAATGA CTACATTTAC ATCTACCCAC ACAACCACCA CTTTGTTCCC TAGTCTTGAT TCTCTCACTC
10151 TAAGTTTCCT GGAGAATCTG AAGTGTATTG GTGGAAGTGG TGCCAAGGAT GAGGGGAGCA ATGAAATATC
10221 TTTCAATAAT ACCACTGCAA CTACTGCTGT TCTTGATCAA TTTGAAGTAT GCTTTGTACA TATTCCATTA
10291 TTTATTTAAT TTCCTTTTTT ATTTGCAATA TTCTATAAAT AATACATTTT ATACCCACTA TACTAAGATA
10361 ATAATTACCT AGAGGGATGG ATGCTATGAC ACAGCTGCTA CACTTCAGAA ACTCTARTAA GGGCAGTTAT
10431 GGAAGTTCAA TAAAATGATA ATGGCATCTT TTGATGGGTA ATATAGGCAA TTTAAGTTTT ATTTCTGTTA
10501 AAGCAGTATT TAGCAAGTAC TGGCCAGTAG GAGAGGAGAA TATCACCTTT TGTGAAAATC TGGTCATTGT
10571 ACCCAGAATT TAGTTAAATG TAACATTTTA GATATTAGGG GTTATCAGGT GACAGATATT GTAGAATAGA
10641 ACAATATGTA ATATTACCCA AAACTATTTT TTCTAAGGTT GCTCTGTTAA ATATGTGCTT TCTTGATTTC
10711 ATTGAATTTG CATTCCTATA TTTTAGGTGG TAAAGTGATT GTCTCTTCAA TAAATCCCGA AATTTTTTAA
10781 TTAAAAAAAA AAAAAACAAA AGTAAATTTT TGATATGGAG AGCACTGGTA TCATTTAGTA TATAAAAAAC
10851 AGATTTTGAA TTAAGTTTCT TATATAAAAG CTGTGTATAT AGTTTAATTA GTTTTACATC ATTTTTCCAT
10921 GTGGTGTTGC AGTTGTCTGA AGCAGGTGGT GTTTCTTGGA GCTTATGCCA ATACGCTAGA GAGATAAAAA
10991 TAGGCAACTG CCATGCATTG TCAAGTGTGA TTCCATGTTA TGCAGCAGTA CAAATGCAGA AAGCTT

SEQ ID NO:23

RLG2B a.a.

MSDPTGIAGAIINPIAQTALVPVTDHVGYMISCRKYVRVMQMKMTELNTSRISVEEHISRNTRNHLQIP
SQTKEWLDQVEGIRANVENFPIDVITCCSLRIRHKLGQKAFKITEQIESLTRQLSLISWTDDPV?LGRVG
SMNASTSASLSDDFPSREKTFTQALIALEPNQKFHMVALCGMGGVGKTRMMQRLKKA?EEKKLFNYIV
GAVI?EKTDPFAIQEAIADYLGIQLNEKTKPARADKLREWFKKNSDGGKTKFLIVLDDVWQLVDLEDIGL
SPFPNQGVDFKVLLTSRDSQVCTMMGVEANSIINVGLLTEAEAQSLFQQFVETSEPELQKIGEDIVRKC
CGLPIAIKTMAC?LRNKRKDAWKDALSRIEHYDIHNVAPKVFETSYHNLQEEETKSTFLMCGLFPEDFDI
PTEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQTNLLIESDDVGCVKMHDLVRAFVLGMFSEVEH
ASIVNHGNMPGWPDENDMIVHSCKRISLTCKGMIEIPVDLKFPKLTILKLMHGDKSLRFPQDFYEGMEKL
HVISYDKMKYPLLPLAPRCSTNIRVLHLTECSLKMFDCSSIGNLSNLEVLSFANSHIEWLPSTVRNLKKL
RLLDLRFCDGLRIEQGVLKSFVKLEEFYIGDASGFIDDNCNEMAERSYNLSALEFAFFNNKAEVKNMSFE
NLERFKISVGCSFDENINMSSHSYENMLQLVTNKGDVLDSKLNGLFLKTEVLFLSVHGMNDLEDVEVKS
THPTQSSSFCNLKVLIISKCVELRYLFKLNLANTLSRLEHLEVCECENMEELIHTGIGGCGEETITFPKLKF
LSLSQLPKLSSLCHNVNIIGLPHLVDLILKGIPGFTVIYPQNKLRTSSLLKEGVVIPKLETLQIDDMENLEE
IWPCELSGGEKVKLRAIKVSSCDKLVNLFPRNPMSLLHHLEELTVENCGSIESLFNIDLDCVGAIGEEDN
KSLLRSINVENLGKLREVWRIKGADNSDLINGFQAVESIKIEKCKRFRNIFTPITANFYLEALLEIQIEGCG
GNHESEEQVTLSISLS

SEQ ID NO:24

SEQ ID NO:

```
                  ---GGAAGAC-ACAATGATGCA-AGGCTGAAGAAGGTTGTGGAAGAAAAGAAAATGTTTAATTATATTGTTGAGGCGGTTATAGGGGAAAAGACAGACCC  → 25
                          10        20        30        40        50        60        70        80        90       100
RLG2A             -GGGGAAGACGACAATGATGCATCGGCTCAAAAAGGTTGTGAAAGAAAAGAAAATGTTTAATTTTATAATTGAGGCGGTTGTAGGGGAAAAAACAGACCC  99 → 27
RLG2B             TGGGGAAGACTAGAATGATGCAAAGGCTGAAGAAGGCTGMTGAAGAAAAGAAATTGTTTAATTATATTGTTGGGGCAGTTATAKGGGAAAAGACGGACCC  100 - 28
RLG2C             -----AACAC--ACGTNCNGCAAAGGCTGAAGAAG-TTGCCAAAGNAAAAGAGGAATTTGGTTATATCATCGAGGCGGTTATAGGGGAAATATCGGACCC  92 - 29
RLG2D             ----------------------------AAGAAG-TTGCCAAAGANAAGAGGAAGTTTGGTTATATCATCGAGGCAGTTATANGGGAAATATCGGACCC  70 - 30
RLG2E             ---GGAAGAC-ACAATGATGCAAAGGTTGAAGAAG-TTGCTAAAGAAAATAGAATGTTCAATTATATGGTTGAGGCAGTTATAGGGGAAAAGACAGACCC  95 - 31
RLG2F             CTTTGGAAGACACAATGATGCAAAGGCTGAAAAAGGTTGTGCATGAAAAGAAAATGTTTAACTTTATTGTTGAAGCAGTTATAGGGGAAAAGACAGACCC  100 - 32
RLG2G             ---GGAAGAC-ACGATGATGAAGAACTG-AAGGAGGTCGTGGACAAAAGAAATCATTCAATATTATTATTCAAGTGGTCATAGGAGAGAAGACAAACCC  95 - 33
RLG2H             --------------------------TG-AAGGAGGTTGTGGAACGAAAGAAAATGTTCAGTATTATTGTTCAAGTGGTCATAGGAGAGAAGACAAACCC  73 - 34
RLG2I             --------AT-GCAA----GAAGAGCTGGAAGGAGGTTGTGGAACAAAAGAAAACGTTCAATATTATTGTTCAAGTGGTCATAGGAGAGAAGACAAACCC  87 - 35
RLG2J             ---GAAAGA------------------------GGTCGTGG-ACANAAGAAAACGTTCAATATTATTGTTCAAGTGGTCATAGGAGAGAAGACAAACCC  71 - 36
RLG2K             TTTGGAAGAC-ACGATGATGAANAGGCTGAAAAATATTATTAAAGAAAAGAGGACGTTTCATTATATTGTTTTGGTGGTTATAAAGGAAAATATGGATCT  99 - 37
RLG2L             ---G----------------------------------------------------TTCAGTTACATGGTCGAGGCAGTTATAGGGGAAAAGACAGACCC  45 - 38
RLG2M             ------------------------GCTGAAGAAG-CTGCTGAAGAAAAGAAATTGTTTAATTATATTGTTGGGGCAGTTATAGGGGAAAAGACGGACCC  74 - 39

                  CATTGCTATTCAGCAAGCTGTAGCAGATTACCTCTGTATAGAGCTGAAAGAAAACACTAAACAAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAGGCC
                         110       120       130       140       150       160       170       180       190       200
RLG2A             CATTGCTATTCAATCAGCTGTAGCAGATTACCTAGGTATAGAGCTCAATGAAAAAACTAAACCAGCAAGAACTGAGAAGCTTCGGAAATGGTTTGTGGAC  199
RLG2B             CTTTGCCATTCAAGAAGCTATAGCAGATTACCTCGGTATACAACTCAATGAAAAAACTAAGCCAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAAAAG  200
RLG2C             CATTGCTATTCAGCAAGTTGTAGCAGATTACCTATGCATAGAACTGAAAGAAAGCGATAAGAAAACAAGAGCTGAGAAGCTTCGTCAAGGGTTCAAGGCC  192
RLG2D             CATTGCTATTCAGCAAGTTGTAGCAGATTACCTATGCATAGAGCTGAAAGAAAGCGATANGAAAACAAGAGCTGAGAAGCTTCGTCAAGGGTTCAAGGCC  170
RLG2E             ACTTGCTATTCAACAAGCTGTAGCGGATTACCTTTGTATAGAGTTAAAAGAAAGCACTAAACCAGCAAGAGCTGATAAGCTTCGTGAATGGTTTAAGGCC  195
RLG2F             CGTTGCCATTCAGGATGCTATAGCAGATTACCTAGGTGTAGAGCTCAATGAAAAATCTAAGCAAGCAAGAGCTGATAAGCTCCGTCAAGGATTCAAGGAC  200
RLG2G             TATTGCAATTCAGCAAGCTGTAGCAGATTACCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGCAAGAGCTGATAAGCTTCGTAAACGGTTTGAAGCC  195
RLG2H             TATTGCTATTCAGCAAGCTGTAGCAGATTACCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGCAAGAGCTGATAAGCTTCGTAAATGGTTCGAGGCC  173
RLG2I             TATTGCTATTCAGCAAGCTGTAGCAGATTCCCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGCAAGAGCTGATAAGCTTCGTAAATGGTTCGAGGCT  187
RLG2J             TATTGCTATCCAGCAAGCTGTAGCAGATTACCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGCAAGAGCTGATAAGCTTCGTAAACGGTTCGAAGCC  171
RLG2K             CATTTCCATCCAGGATGCTGTAGCAGATTATCTGGATATGAAGCTAACAGAAAGCAATGAATCAGAAAGAGCCGATAAACTTCGTGAAGGGTTTCAGGCC  199
RLG2L             CATTGCTATTCAACAAGCTGTANCCGATTACCTNCGTATACAGTTCAAAGAAAGCACTAAACCAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAGGCC  145
RLG2M             CTTTGCCATTCAAGAAGCTATAGCAGATTACCTCGGTATACAACTCAATGAAAAAACTAAGCCAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAAAAG  174
```

```
       AAATCAGATGGAGGTAAGAATAAGTTCCTCGTAATACTTGACGATGTTTGGCAGTTTGTTGATCTTGAAGATATTGGTTTAAGTCCTTTTCCAAATCAAG
               210       220       230       240       250       260       270       280       290       300
RLG2A  AATTCTGGTG---GTAAGA---AGATCCTAGTCATACTCGACGATGTATGGCAGTTTGTGGATCTGAATGATATTGGTTTAAGTCCTTTACCAAATCAAG 293
RLG2B  AATTCAGATGGAGGTAAGACTAAGTTCCTCATAGTACTTGACGATGTTTGGCAATTAGTTGATCTTGAAGATATTGGGTTAAGTCCTTTTCCAAATCAAG 300
RLG2C  AAATCAGATGGAGGTAACACTAAGTTCCTCATAATATTGGATGATGTCTGGCAGTCCGTTGATCTANAAGATATTGGTTTAAGCCCTTCTCCCAATCAAN 292
RLG2D  AAATCANATGGAGGTAACACTAAGTTCCTCATAATATTGGATGATGTCTGGCAGTCCGTTGATCTAGAANATATTGGTTTAAGCCCTTCTCCCAATCAAG 270
RLG2E  AACTCTGGAGAAGGTAAGAATAAGTTCCTTGTAATATTTGATGATGTTTGGCAGTCCGTTGATCTGGAAGACATTGGTTTAAGTCATTTTCCAAATCAAG 295
RLG2F  AAATCAGATGGAGGCAAAAATAAGTTCTTTGTAATACTTGACGATGTTTGGCAGTCTGTTGATCTGGAAGATATTGGTTTAAGTCCTTTTCCAAATCAAG 300
RLG2G  ------GATGGAGGAAAGAATAAGTTCCTTGTAATACTTGACGATGTATGGCAGTTTGTCGATCTTGAAGATATTGGTTTAAGTCCTCTGCCAAATAAAG 289
RLG2H  ------GATGGAGGAAAGAATAAGTTCCTTGTAATACTTGACGATGTATGGCAGTTTGTCGATCTTGAAGATATTGGTTTAAGTCCTCTGCCAAATAAAG 267
RLG2I  ------GATGGAGGAAANAATAAGTTCCTCGTNATACTTGACGATGTATGGCNGTTTGTTGATCTTGAANATATTGGTTTAAGTCCTCATCCAAATNAAG 281
RLG2J  ------GATGGAGGAAAGAATAAGTTCCTTGTNATACTTGACGATGTTTGGCAGTTTTTCGATCTTGAANATATTGGTTTAAGTCCTCTGCCAAATAAAG 265
RLG2K  AAATCAGATGGAGGTAAGAATAGGTTCCTCATAATACTGGATGATGTATGGCAATCTGTTAATATGGAAGATATTGGTTTAAGTCCTTTTCCGAATCAAG 299
RLG2L  NACTCTGNAGANGGTAAGAATAAGTTCCTNGTAATATTTGATGACGTCTGGCAGTCCGTTGATCTGGAAGATATTGGNTTAAGTCCTTTTCCAAATCAAN 245
RLG2M  AATTCAGATGGAGGTAAGACTAAGTTCCTCATANTACTTGACGATGTTTGGCAATTAGTTGATCTTGAAGATATTGGGTTAAGTCCTTTTCCAAATCAAG 274

       GTGTCGACTTCAAGGTCTTGTTGACTTCACGAGACTCACATGTTTGCACTGTGATGGGGGTTGAAGCTAATTCAATTCTTAATGTGGACTTCTAATAGA
               310       320       330       340       350       360       370       380       390       400
RLG2A  GTGTCGACTTCAAGGTGTTGTTGACATCACGAGACAAAGATGTTTGCACTGAGATGGGAGCTGAAGTTAATTCAACTTTTAATGTGAAAATGTTAATAGA 393
RLG2B  GTGTCGACTTCAAGGTCTTGTTGACATCACGAGACTCACAAGTTTGCACTATGATGGGGGTTGAAGCTAATTCAATTATTAACGTGGGCCTTCTAACTGA 400
RLG2C  GTGTCGACTTCAAGGTCTTGTTGACTTCACGAGACGAACATGTTTGCTCAGTGATGGGGGTTGAAGCTAATTCAATTATTAACGTGGGACTTCTAATTGA 392
RLG2D  GTGTCNACTTCAAGGTCTTGTTGACTTCACGAGACGAACATGTTTGCTCAGTGATGGGGGTTGAAGCTAATTCAATTATTAACGTGGGACTTCTAATTGA 370
RLG2E  GTGTCGACTTCAAGGTCTTGTTGACTTCACGAGACTAACATGTKTGCMCAGTAATGGGGGTTGAAGCTAATTCAATTCTTAATGTGGGACTTCTAGTAGA 395
RLG2F  GCGTCGACTTCAAGGTCTTGTTGACATCACGAGACAGACATGTTTGCACAGTGATGGGGGTTGAAGCCAAATTAATTCTAAACGTGGGACTTCTAATTGA 400
RLG2G  GTGTCAACTTCAAGGTCTTGTTGACGTCAAGAGATTCACATGTTTGCACTCTGATGGGAGCTGAAGCAAATTCAATTCTTAATATAAAAGTTTTAAAAGA 389
RLG2H  GTGTCAACTTCAAGGTCTTGTTGACGTCAAGAGATTCACATGTTTGCACTCTGATGGGAGCCGAAGCCAATTCAATTCTCAATATAAAAGTTTTAACAGC 367
RLG2I  GTGTCANCTTCAAGGTCTTGTTGACGTCAAGAGATTCACATGTTTGCACTCTGATGGGAGCTGAAGCCAATTCAATTCTCAATATAAAAGTTTTAAAAGA 381
RLG2J  GTGTCNACTTCAAGGTCTTGTTGACGTCNAGANATTCNCATGTTTGCACTCTGATGGGAGCTGAAGCCAATTCNATTCTCAATNTAAAAGTTTTANAAGA 365
RLG2K  GTGTCGACTTCAAGGTCTTGTTGACCTCGGAAAACAAAGATGTTTGTGCAAAAATGGGAGTTGAAGCTAATTTAATTTTCGACGTGAAATTCTTAACAGA 399
RLG2L  GTGTCGACTTCAAGGTCTTGTTGACTTCACGAGACNAACACGTTTGCACAATGATGGGGGTTNAAGCTAATTCANTTATTNATGTGGGACTTCTAACTGA 345
RLG2M  GTGTCGACTTCAAGGTCTTGTTGACATCACGAGACTCACAAGTTTGCACTATGATGGGGGTTGAAGCTAATTCAATTATTAACGTGGGCCTTCTANCTGA 374
```

```
            AGCAGAAGCACAAAGTTTGTTCCACCAATTTGTAGAAACTTC--------------TGAGCCCGAGCTCCA-TAAGATAGGAGAAGATATTGTAAGGAAG
                     410       420       430       440       450       460       470       480       490       500
RLG2A       AACAGAAGCACAAAGTTTATTCCACCAATTTATAGAAATTTCGGATGATG-----TTGATCCTGAGCTCCA-TAATATAGGAGTGAATATTGTAAGGAAG 487
RLG2B       AGCAGAAGCTCAAAGTCTGTTCCAACAATTTGTAGAAACTTC--------------TGAGCCCGAGCTCCA-GAAGATAGGAGAGGATATCGTAAGGAAG 485
RLG2C       AGCAGAAGCACAAAGATTGTTCCAGCAATTTGTAGAAACTTC--------------TGAGCCCGAGCTCCA-CAAGATAGGAGAAGATATTGTTAGGAGG 477
RLG2D       ANCANAANCACNAAGATTGTTCCAGCAATTTGTANAAACTTC--------------TGAGCCCGAGCTCCA-CAAGATANGAGAANATATTGTTAGGAGG 455
RLG2E       AGCAGAAGCACAAAGTTTGTTCCAGCAATTTGTAGAAACTTT--------------TGAGCCCGAGCTCCA-TAAGATAGGAGAAGATATCGTAAGGAAG 480
RLG2F       AGCTGAAGCACAAAGTTTGTTCCACCAATTTGTTGTCACTTC--------------TGAGCCCGAGCTCCA-TAAGATAGGAGAAGATATTGTAAAGAAG 485
RLG2G       TGTAGAAGGACAAAGTTTGTTCCGCCAGTTTGCTAAAAAATGCGGGTGATGATGACCTGGATCCTGCTTTCAATGGGATAGCAGATAGTATTGCAAGTAGA 489
RLG2H       TGTAGAAGGACAAAGTTTGTTCCGCCAGTTTGCTAAAAAATGCGGGTGATGATGACCTGGATCCTGCTTTCAATAGGATAGCAGATAGTATTGCAAGTAGA 467
RLG2I       TGTAGAAGGAAAAAGTTTGTTCCGCCAGTTTGCTAAAAAATGCGGGTGATGATGACCTGGATCCTGCTTTCATTGGGATAGCAGATAGTATTGCAAGTAGA 481
RLG2J       TGTTCAAGGAAAAAGTTTGT-CCGCCANTTTGCTAAAAAATGCGGGTGATNATGACCTGGATCCNGCTTTCATTGGGATANCANATAGTNTTGCCAGTNGA 464
RLG2K       AGAAGAAGCACAAAGTTTGTTTTATCAATTTGTAAAAGTTTCTGATA--------------CCCACCTTGA-TAAGATTGGAAAAGCTATTGTAAGAAAC 484
RLG2L       AGTAGAAGCACNAAGTCTGTTCCANCAATTTGTAGAAACTTN--------------TGAGCCCGAGCTCTG-TAAGATANGANAAGTTATCGTAAGAAAG 430
RLG2M       AGCAGAAGCTCAAAGTCTGTTCCAACAATTTGTAGAAACTTC--------------TGAGCCCGAGCTCCA-GAAGATAGGAGAGGATATCGTAAGGAAG 459

            TGTTG-CGGTCTACCCATTGCCATCAAAACCATGGCCTGTACTCTTAGAAATAAAAGAAAGGATGCATGGAAGGATGCACTTTCTCGTTTAGAGCACCAT
                     510       520       530       540       550       560       570       580       590       600
RLG2A       TGTGG-GGGTCTACCCATTGCCATAAAAACCATGGCGTGTACTCTTAGAGGAAAAAGCAAGGATGCATGGAAGAATGCACTTCTTCGTTTAGAGCACTAT 586
RLG2B       TGTTG-CGGTCTACCTATTGCCATAAAAACCATGGCATGTWCTCTTAGAAATAAAAGAAAGGATGCATGGAAGGATGCACTTTCGCGCATAGAGCACTAT 584
RLG2C       TGTTG-CGGTCTACCCATTGCCATCAAAACCATGGCGTGTACTCTAAGAAATAAAAGAAAGGATGCATGGAAGGATGCACTTTCTCGTTTACAACACCAT 576
RLG2D       TGTTG-CGGTCTACCCATTGCCATCAAAACCATGGCGTGTACTCTAAGAAATAAAAGAAAGGATGCATGGAAGGATGCACTTTCTCGTTTACAACACCAT 554
RLG2E       TGTTG-TGGTTTACCTATTGCCATTAAAACCATGGCATGTACTCTAAGAAATAAAAGAAAGGATGCATGGAAGGATGCACTTTTGCATTTAGAGTACCAT 579
RLG2F       TGTTT-CGGTCTGCCAATTGCCATCAAAACCATGGCATGTACTCTACGACATAAAAGAAAGGATGCATGGAAGGATGCACTTTCACGTTTAGAGCACCAT 584
RLG2G       TGTCAAGG-TTTGCCCATTGCCATCAAAACCATTGCCTTAAGTCTTAAAGGTAGAAGCAAGTCTGCATGGGACGTTGCACTTTCTCGTCTGGAGAATCAT 588
RLG2H       TGTCAAGG-TTTGCCCATTGCCATCAAAACCATTGCCTTAAGTCTTAAAGGTAGAAGCAAGCCTGCGTGGGACCATGCGCTTTCTCGTTTGGAGAACCAT 566
RLG2I       TGTCAAGG-TTTGCCCATTGCCATCAAAACCATTGCCTTAAGTCTTAAAGGTAGAAGCAAGTCTGCATGGGACGTTGCACTTTCTCGTCTGGAGAATCAT 580
RLG2J       TGTCNAGGGTTTGNCCATTGCCNTCAAAACCATTGNCTTNAGTCTTAAAGGTAGAAGCAAGTCTGCATGGGACGTCGCACTTTCTCGTCTGGAGAATCAT 564
RLG2K       TGTGG-TGGTCTACCCATTGCCATCAAAACCATAGCCAATACTCTTAAAAATANAAACAAGGATGTTTGGAAGGATGCACTTTCTCGTATAGAGCATCAT 583
RLG2L       TGTTG-CGGTCTACCTATTGCCATCAAAACCATGGCGTGTTCTCTTAGAAATAAAAGAAAGGATGCATGGAAGGATGCACTTTCNCGTATAGAGCACTAT 529
RLG2M       TGTTG-CGGTCTACCTATTGCCATAAAAACCATGGCATGTACTCTTAGAAATAAAAGAAAGGATGCATGGAAGGATGCACTTTCGCGCATAGAGCACTAT 558
```

```
       GACATTGGTAGTG------TTGCGCCTGAAGTTTTTAAAACGAGCTACGACAATCTCCAAGACGAGGAGACTAAATCTATTTTTTTGATGTGTGGTTTGT
                 610       620       630       640       650       660       670       680       690       700
RLG2A  GACATTGAAAATA------TTGTTAATGGAGTTTTTAAAATGAGTTACGACAATCTCCAAGATGAGGAGACTAAATCCACCTTTTTGCTTTGTGGAATGT 680
RLG2B  GACATTCACAATG------TTGCGCCCAAAGTCTTTGAAACGAGCTACCACAATCTCCAAGAAGAGGAGACTAAATCCACTTTTTTAATGTGTGGTTTGT 678
RLG2C  GACATTGGTAATG------TTGCTACTGCAGTTTTTAGAACCAGCTATGAGAATCTCCCGGACAAGGAGACAAAATCTGTTTTTTTGATGTGTGGTTTGT 670
RLG2D  GACATTGGTAATG------TTGCTACTGCAGTTTTTANAACCAGCTATGAGAATCTCCCGGACAAGGAGACAAAATCTGTTTTTTTGATGTGTGGTTTGT 648
RLG2E  GACATTAGCAGTG------TTGCGCCCAAAGTCTTTGAAACGAGCTACCATAATCTCCACAACAAGGAGACTAAATCTGTGTTTTTGATGTGTGGTTTTT 673
RLG2F  GACATTCAAAGTG------TTGTGCCTAAAGTATTTGAAACGAGCTACAACAATCTCAAAGACAAGGAGACTAAATCCGTATTTTTGATGTGTGGTTTGT. 678
RLG2G  AAGATTGGTAGTGAAGAAGTTGTGCGTGAAGTTTTTAAAATTAGCTACGACAATCTCCAAGATGAGGTTACTAAATCTATTTTTTTACTTTGTGCTTTAT 688
RLG2H  AAGATTGGTAGTGAAGAAGTTGTGCGTGAAGTTTTTAAAATTAGCTATGACAATCTCCAAGATGAGATTACTAAATCTATTTTTTTACTTTGTGCTTTAT 666
RLG2I  AAGATTGGTAGTGAAGAAGTTGTGCGTGAAGTTTTTAAAATTAGCTATGACAATCTCCAAGATGAGGTTACTAAATCTATTTTTTTACTTTGTGCTTTAT 680
RLG2J  AAGATTGGTAGTGAAGAAGTTGTGCGNGAAGTTTTTAAAATNAGCTATGACAATCTCCAAGATGAGGTTACTAAATCTATTTTTTTACTCTGTGCTTTAT 664
RLG2K  GACATTGAGACAA------TTGCACATGTTGTTTTTCAAATGAGCTACGACAATCTCCAAAACGAAGAAGCTCAATCCATTTTTTTGCTTTGTGGATTGT 677
RLG2L  GACATTCNTNGTG------TTGCGCCCAAAGTCTTTGAAACAAAAAACCATAATCTCCACAACAAGGAGACTAAATCCGCGTTTTTGATGTGTGGTTTGT 623
RLG2M  GACATTCACAATG------TTGCGCCCAAAGTCTTTGAAACGAGCTACCACAATCTCCAAGAAGAGGAGACTAAATCCACTTTTTTAATGTGTGGTTTGT 652

       TTCCTGAAGACTTTGATATTCCTACTGAGGAGTTGATGAGGTATGGATGGGGCTTGAAATTATTTGATAGAGTTTATACTATTAGAGAAGCAAGAAACAG
                710       720       730       740       750       760       770       780       790       800
RLG2A  ATCCCGAAGACTTTGATATTCTTACCGAGGAGTTGGTGAGGTATGGATGGGGGTTGAAATTATTTAAAAAAGTGTATACTATAGGAGAAGCAAGAACCAG 780
RLG2B  TTCCCGAAGACTTCGATATTCCTACTGAGGAGTTGATGAGGTATGGATGGGGCTTGAAGCTATTTGATAGAGTTTATACGATTAGAGAAGCAAGAACCAG 778
RLG2C  TTCCCGAAGACTTCAATATTCCTACCGAGGAGTTGATGAGGTATGGATGGGGCTTAAAGTTATTTGATAGAGTTTATACAATTATAGAAGCAAGAAACAG 770
RLG2D  TTCCCGAAGACTTCAATATTCCTACCGAGGAGTTGATGANGTATGGATGGGGCTTAAAGTTATTTGATAGAGTTTATACAATTATAGAAGCAAGAAACAG 748
RLG2E  TTCCTGAAGACTTCAATATTCCAATCGAGGAGTTGATGAGGTATGGATGGGGCTTAAAGATATTTGATAGAGTTTATACTATTAGACAAGCAAGAATCAG 773
RLG2F  TTCCTGAAGACTTGGATATACCTATCGAGGAGTTGATGAGGTATGGATGGGGCTTAAGATTATTTGATAGAGTTAATACTATTACACAAGCAAGAAACAG 778
RLG2G  TTCCTGAAGATTTTGATATTCCTACTGAGGAGTTGGTGAGGTATGGGTGGGGCTTGAAATTATTTATAGAAGCAAAAACTATAAGAGAAGCAAGAAACAG 788
RLG2H  TTCCTGAAGATTTTGATATTCCTACTGAGGAGTTGATGAGGTATGGATGGGGCTTGAAATTATTTATAGAAGCAAAAACTATAAGAGAAGCAAGAAACAG 766
RLG2I  TTCCTGAAGATTTTGATATTCCTACTGAGGAGTTGGTGAGGTATGGGTGGGGCTTGAAATTATTTATAGAAGCAAAAACTATAAGAGAAGCAAGAAACAG 780
RLG2J  TTCCTGAAGATTTTGATATTCCTATTGAGGAGTTGGTGAGGTATGGGTGGGGCTTGAAATTATTTATAGAAGCAAAAACTATAAGAGAAGCAAGAAACAG 764
RLG2K  TTCCTGAAGACTTTGATATTCCTACTGAGGAATTGGTGANGTATGGATGGGGATTGAGAGTATTTAATGGAGTGTATACTATAGGAGAAGCAAGACACAG 777
RLG2L  TTCCTGAAGACTTCAATATTCCTACCGAGGAGTTGATGAGGTATGGATGGGGCTTAAAGCTATTTGACAGAGTTTATACAATTAGAGAAGCAAGAACCAG 723
RLG2M  TTCCCGAAGACTTCGATATTCCTACTGAGGAGTTGATGAGGTATGGATGGGGCTTGAAGCTATTTGATAGAGTTTATACGATTAGAGAAGCAAGAACCAG 752
```

```
         GCTCAACACCTGCATTGAGCGACTGGTGCAGACAAATTTGTTAATTGAAAGTGATGATGTTGGGTGCGTCAAGATGCATGATCTGGTGCGTGCTTTTGTT
                 810       820       830       840       850       860       870       880       890       900
RLG2A    GCTCAACACATGCATTGAGCGGCTCATTCATACAAATTTGTTGATGGAAGTTGATGATGTTAGGTGCATCAAGATGCATGATCTTGTTCGTGCTTTTGTT 880
RLG2B    GCTCAACACCTGCATTGAGCGACTGGTGCAGACAAATTTGTTAATTGAAAGTGATGATGTTGGGTGTGTCAAGATGCATGATCTGGTCCGTGCTTTTGTT 878
RLG2C    GCTCAACACCTGCATTGACCGACTGGTGCAGACAAATTTACTAATTGGAAGTGATAATGGTGTACATGTCAAGATGCATGATCTGGTCCGTGCTTTTGTT 870
RLG2D    GCTCAACACCTGCATTGAGCGACTGGTGCAGGCAAATTTACTAATTGGAAGTGATAATGGTGTACACGTCAAGATGCATGATCTGGTCCGTGCTTTTGTT 848
RLG2E    GCTCAACACCTGCATTGAGCGACTGGTGCAGACAAATTTGTTAATAGAAAGTGATGATGGTGTGCACGTCAAGATGCATGATCTGGTCCGTGCTTTCGTT 873
RLG2F    GCTCAACACCTGCATTGAGCGACTGGTGCACACAAATTTGTTAATTGAAAGTGTTGATGGTGTGCATGTCAAGATGCATGATCTGGTTCGTGCTTTTGTT 878
RLG2G    GCTCAACACCTGCACTGAGCGGCTTAGGGAGACAAATTTGTTATTTGGAAGTGATGACATTGGATGTGTCAAGATGCACGATGTGGTGCGTGATTTTGTT 888
RLG2H    GCTCAACACCTGCACTGAGCGGCTTAGGGAGACAAATTTGTTATTTGGAAGCGATGACATTGGATGCGTCAAGATGCACGATGTGGTGCGTGATTTTGTT 866
RLG2I    GCTCAACACCTGCACTGAGCGGCTTAGGGAGACAAATTTGTTATTTGGAAGTGATGACATTGGATGCGTCAAGATGCACGATGTGGTGCGTGATTTTGTT 880
RLG2J    GCTCAACAACTGCACTGAGCGGCTTAGGGAGACAAATTTGTTATTTGGAAGTCATGACTTTGGGTGCGTCAAGATGCACGATGTGGTGCGTGATTTTGTT 864
RLG2K    GTTGAACGCCTACATCGAGCTGCTCAAGGATTCTAATTTATTGATTGAAAGTGATGATGTTCACTGCATCAAGATGCATGATTTAGTTCGTGCTTTTGTT 877
RLG2L    GCTCAACACCTGCATTGAGCGACTTGTGCAGACAAATTTGTTAATTGAAAGTGATGATGTTGGGTGTGTCAAGATGCATGATCTGGTGCGTGCTTTTGTT 823
RLG2M    GCTCAACACCTGCATTGAGCGACTGGTGCAGACAAATTTGTTAATTGAAAGTGATGATGTTGGGTGTGTCAAGATGCATGATCTGGTCCGTGCTTTTGTT 852

         TTGGGTATGTTTTCTGAAGTCGAGCATGCTTCAATTGTCAACCATGGTAATA-TGCCTGGGTGGCCTGA---GAAAATGATATGATCGTGTACTCTTGCA
                 910      920       930       940       950       960       970       980       990      1000
RLG2A    TTGGATATGTATTCTAAAGTCGAGCATGCTTCCATTGTCAACCATAGTAATA-CACTAGAGTGGC---A--TGCAGATAATATG--CACG-ACTCTTGTA 971
RLG2B    TTGGGTATGTTTTCTGAAGTCGAGCATGCTTCTATTGTCAACCATGGTAATA-TGCCTGGGTGGCCTGA--TGAAAATGATATGATCGTGCACTCTTGCA 975
RLG2C    TTGGGTATGTATTCTGAAGTCGAGCAAGCTTCAATTGTCAACCATGGTAATA-TGCCTGGGTGGCCTGA--TGAAAATGATATGATCGTGCACTCTTGCA 967
RLG2D    TTGGGTATGTATTCTGAAGTCGAGCAAGCTTCAATTGTCAACCATGGTAATA-TGCCTGGGTGGCCTGA--TGAAAATGATATGATCGTGCACTCTTGCA 945
RLG2E    TTGGTTATGTTTTCTGAAGTTGAACATGCTTCAATTATCAACCATGGTAATA-TGCTTGGATGGCCTG-----AAAATTATATGACCA---ACTCTTGCA 964
RLG2F    TTGGGAATGTTTTCTGAAGTGGAGCATGCTTCAATTGTCAACCATGGTAATA-TGCCCGAGTGGACTG-----AAAATGATATGACTG---ACTCTTGCA 969
RLG2G    TTGCATATATTCTCAGAAGTCCAACACGCTTCAATTGTCAACCATGGTAAC-GTGTCAGAGTGGCTAGAG--GAAAATCAT---AGCATCTACTCTTGTA 982
RLG2H    TTGCATATATTCTCAGAAGTCCAGCACGCTTCAATTGTCAACCATGGTAAC-GTGTCAGAGTGGCTAGAG--GAAAATCAT---AGCATCTACTCTTGTA 960
RLG2I    TTGCATATATTCTCAGAAGTCCAGCACGCTTCAATTGTCAACCATGGTAAT-GTGTCAGAGTGGCTAGAG--GAAAATCAT---AGCATCTACTCTTGTA 974
RLG2J    TTGCATATGTTTTCAGAAGTCAAGCATGCTTCAATTGTCAACCATGGTAAC-ATGTCAGAGTGGCCAGAG--AAAAATGAT---ACCAGCAACTCTTGTA 958
RLG2K    TTGGATACGTTTAATAGATTCAAGCATTCTTTGATTGTTAACCATGGTAATGGTGGTATGTTAGGGTGGCCTGAAAATGATATGAGTGCCTCATCTTGCA 977
RLG2L    TTGGGTATGTATTCTGAAGTCGAGCATGCTTCAATTGTCAACCATGGTAATA-TGCATGGGTGGACTA-----AAAATGATATGAACG---ACTCTTGCA 914
RLG2M    TTGGGTATGTTTTCTGAAGTCGAGCATGCTTCTATTGTCAACCATGGTAATA-TGCCTGGGTGGCCTGA--TGAAAATGATATGATCGTGCACTCTTGCA 949
```

```
      AAAGAATTTCATTAACATGCAAGGGTATGTCTGAGTTTCCAGTAGACCTCAAGTTTCCAAACCTAACGATTTTGAAACTTATGCATGGAGATAAGTCGCT
                1010      1020      1030      1040      1050      1060      1070      1080      1090      1100
RLG2A AAAGACTTTCATTAACATGCAAGGGTATGTCTAAGTTTCCTACAGACCTGAAGTTTCCAAACCTCTCCATTTTGAAACTTATGCATGAAGATATATCATT 1071
RLG2B AAAGAATTTCATTAACATGCAAGGGTATGATTGAGATTCCAGTAGACCTCAAGTTTCCTAAACTAACGATTTTGAAACTTATGCATGGAGATAAGTCGCT 1075
RLG2C AAAGAATTTCATTAACATGCAAGGGTATGATTGAGTTTCCAGTAGACCTCAAGTTTCCTAAACTAACGATTTTGAAACTTATGCATGGAGATAAATCGCT 1067
RLG2D AAAGAATTTCATTAACATGCAAGGGTATGATTGAGATTCCAGTAGACCTCAAGTTTCCTAAACTAACGATTTTGAAACTTATGCATGGAGATAAGTCTCT 1045
RLG2E AAACAATTTCATTAACATGCAAGAGTATGTCTGAATTTCCGGGAGATCTCAAGTTTCCAAACCTAACGATTTTGAAACTCATGCATGGAGATAAGTTGCT 1064
RLG2F AACAAATTTCATTAACATGCAAGAGTATGTTGGAGTTTCCTGGAGACCTCAAGTTTCCAAACCTAAAGATTTTGAAACTTATGCATGGAGGTAAGTCACT 1069
RLG2G AAAGAATTTCATTAACATGCAAGGGTATGTCTCAGTTTCCCAAAGACCTCAAATTTCCAAACCTTTCAATTTTGAAACTTATGCATGGAGATAAGTCACT 1082
RLG2H AAAGAATTTCATTAACATGCAAGGGTATGTCTGAGTTTCCCAAAGACCTCAAATTTCCAAACCTTTCAATTTTGAAACTTATGCATGGAGATAAGTCGCT 1060
RLG2I AAAGAATTTCATTAACATGCAAGGGTATGTCTGAGTTTCCCAAAGACCTCAAATTTCCAAACCTTTCAATTTTGAAACTTATGCATGGAGATAAGTCGCT 1074
RLG2J AAAGAATTTCATTAACATGCAAGGGTATGTCTAAGTTTCCTAAAGACATCAACTATCCAAACCTTTTGATTTTGAAACTTATGCATGGAGATAAGTCGCT 1058
RLG2K AAAGAATTTCATTAATATGCAAGGGCATGTCCGATTTTCCTAGAGACGTAAAGTTTCCAAATCTCTTGATTTTGAAACTTATGCATGCAGATAAGTCTTT 1077
RLG2L AAACAGTTTCTTTAACATGCGAGAGTGTGTCTGAGTTTCCAGGAGACCTCAAGTTTCCAAACCTAAAGCTTTTGAAACTTATGCATGGAGATAAGATGCT 1014
RLG2M AAAGAATTTCATTAACATGCAAGGGTATGATTGAGATTCCAGTAGACCTCAAGTTTCCTAAACTAACGATTTTGAAACTTATGCATGGAGATAAGTCGCT 1049

      AAGGTTTCCTCAAGACTTTTATGAAGGAATGGAAAAGCTTCAGGTTATATCATACGATAAAATGAAGTATCCATTGCTTCCCTCGTCACCTCAATGCTCC
               1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
RLG2A GAGGTTTCCCAAAAACTTTTATGAAGAAATGGAGAAGCTTGAGGTTATATCCTATGATAAAATGAAATATCCATTGCTTCCCTCATCACCTCAATGTTCC 1171
RLG2B AAGGTTTCCTCAAGACTTTTATGAAGGAATGGAAAAGCTCCATGTTATATCATACGATAAAATGAAGTACCCATTGCTTCCTTTGGCACCTCGATGCTCC 1175
RLG2C AAAGTTTCCTCAAGAATTTTATGAAGGAATGGAAAAGCTCCGGGTTATATCATACCATAAAATGAAGTACCCATTGCTTCCTTTGGCACCTCAATGCTCC 1167
RLG2D AAAGTTTCCTCAAGAATTTTATGAAGGAATGGAAAAGCTCCAGGTTATATCATACGATAAAATGAAGTACCCATTGCTTCCTTTGGCACCTCAATGCTCC 1145
RLG2E AAGATATCCTCAAGACTTTTATGAAGGAATGGAAAAGCTCTGGGTTATATCATATGATGAAATGAAGTATCCATTGCTTCCCTCGTTACCTCAATGCTCC 1164
RLG2F AAGGTATCCTCAAGACTTTTATCAAGGAATGGAAAAGCTGGAGGTTATATCATACGATGAAATGAAGTATCCATTGCTTCCCTCGTTGCCTCAATGTTCC 1169
RLG2G GAGCTTTCCTGAAAACTTTTATGGAAAGATGGAAAAGGTTCAGGTAATATCATATGATAAATTGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCC 1182
RLG2H GAGCTTTCCTGAAAACTTTTATGGAAAGATGGAAAAGGTTCAGGTAATATCATATGATAAATTGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCC 1160
RLG2I GAGCTTTCCTGAAAACTTTTATGGAAAGATGGAAAAGGTTCAGGTAATATCATATGATAAATTGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCC 1174
RLG2J GTGCTTTCCTGAAAACTTTTATGGAAAGATGGAAAAGGTTCAGGTAATATCATATGATAAATTGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCC 1158
RLG2K GAAGTTTCCTCAAGACTTTTATGGAGAAATGAAGAAGCTTCAGGTTATATCATACGATCACATGAAGTATCCCTTGCTTCCAACATCACCTCAATGCTCC 1177
RLG2L AAGGTTTTCTCAAGACTTTTATGAAGGAATGGAAAAGCTCCAGGTAATATCATACCATAAAATGAAGTATCCATTGCTTCCCTCGTCACCTCAATGCTCC 1114
RLG2M AAGGTNTCCNCNAGACTTTTATGAAGGAATGGAAAAGCTCCNTGTTATATCATACGATAAAATGAAGTNCCCATTGCTTCCTTTGGCACCTCGATGCTCC 1149
```

```
       ACCAACCTTCGAGTGCTTCATCTCCATGAATGTTCATTAAGGATGTTTGATTGCTCTTCTATTGGAAATCTTTTGAATCTGGAAGTGCTCAGCTTTGCTA
                1210      1220      1230      1240      1250      1260      1270      1280      1290      1300
RLG2A  GTCAACCTTCGCGTGTTTCATCTACATAAATGCTCGTTAGTGATGTTTGACTGCTCTTGTATTGGAAATCTGTCGAATCTAGAAGTGCTTAGCTTTGCTG 1271
RLG2B  ACCAACATTCGGGTGCTTCATCTCACTGAATGTTCATTAAAGATGTTTGATTGCTCTTCTATCGGAAATCTATCGAATCTGGAAGTGCTGAGCTTTGCAA 1275
RLG2C  ACCAACATTCGGGTGCTTCATCTCACGGAATGTTCATTAAAGATGTTTGATTGCTCGTGTATTGGAAATCTATCGAATCTGGAAGTGCTGAGCTTTGCTA 1267
RLG2D  ACCAACATTCGGGTGCTTCATCTCACTGAATGTTCATTAAAGATGTTTGATTGCTCTTCTATCGGAAATCTATCGAATCTGGAAGTGCTGAGCTTTGCTA 1245
RLG2E  ATCAACCTTCGAGTGCTTCACCTCCATCGATGCTCATTAATGATGTTTGATTGCTCTTGTATTGGAAATATGTTGAATCTGGAAGTGCTTAGCTTTGTTA 1264
RLG2F  ACCATCCTTCGAGTGCTTCATCTCCATGAATGTTCATTAAGGATGTTTGATTGCTCTTCAATCGGTAATCTTTTCAACATGGAAGTGCTCAGCTTTGCTA 1269
RLG2G  ACCAACGTTCGAGTGCTTCATCTTCATTACTGTTCATTAAGGATGTTTGATTGCTCTTCAATTGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTA 1282
RLG2H  ACTAACGTTCGAGTGCTTCATCTCCATTATTGTTCATTAAGGATGTTTGATTGCTCTTCAATTGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTA 1260
RLG2I  ACCAACCTTCGAGTGCTTCATCTCCATGAATGTTCATTAAGGATGTTTGATTGCTCTTCAATTGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTA 1274
RLG2J  ACTAACGTTCGAGTGCTTCATCTCCATTATTGTTCATTAAGGATGTTTGATTGCTCTTCAATTGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTA 1258
RLG2K  ACCAACCTTCGTGTGCTTCATCTTCATCAATGCTCATT---GATGTTTGATTGCTCTTCTATTGGAAATCTGTTGAATCTGGAAGTGCTCAGCTTTGCTA 1274
RLG2L  ACCAACCTTCGAGTGCTTCATCTTCATCGGTGTTCATTACGGATGCTTGATTGCTCTTGTATCGGAAATTTGACGAATCTGGAAGTGTTGAGCTTCGCTA 1214
RLG2M  ACCAACATTCGGGTGCTTCATCTCACTGAATGTTCATTAAAGATGTTTGATTGCTCNTCNATCGGAAATCTATCNAATCTGGAAGTGCTGANCTTTGCAA 1249

       ATTCTGGCATTGAATGGTTACCTTCCACAATCGGAAATTTGAAGAAGCTAAGGCTACTTGATTTGACAAATTGTTATGGTCTTCGTATAGAAAATGGTGT
                1310      1320      1330      1340      1350      1360      1370      1380      1390      1400
RLG2A  ATTCTGCCATTGACCGGTTGCCTTCCACAATCGGAAAGTTGAAGAAGCTAAGGCTACTGGATTTGACGAATTGTTATGGTGTTCGTATAGATAATGGTGT 1371
RLG2B  ATTCTCACATTGAATGGTTACCTTCCACAGTCAGAAATTTAAAGAAGCTAAGGTTACTTGATCTGAGATTTTGTGATGGTCTCCGTATAGAACAGGGTGT 1375
RLG2C  ATTCTTGCATTGAGTGGTTACCTTCCACGGTCAGAAATTTAAAAAAGCTAAGGTTACTTGATTTGAGATTGTGTTATGGTCTCCGTATAGAACAGGGTGT 1367
RLG2D  ATTCTCGCATTGAATGGTTACCTTCCACAGTCAGAAATTTAAAGAAGCTAAGGTTACTTGATCTGAGATTTTGTGATGGTCTCCGTATAGAACAGGGTGT 1345
RLG2E  AATCTGGCATTGAATGGTTACCTTCCACAATAGGAAATTTAAAGAAGCTAAGGTTACTTGATCTGAGAGATTGTTATGGTCTTCGTATAGAAAAAGGTGT 1364
RLG2F  ATTCTAGCATTGAATTGTTACCTTCCGTAATTGGAAATTTGAAGAAGTTGCGGCTGCTAGATTTGACAAACTGTTATGGTGTTCGTATAGAAAAGGATGT 1369
RLG2G  ATTCTAACATTGAATGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGCTACTAGATTTGACAAATTGTAAAGGTCTTCGTATAGATAATGGTGT 1382
RLG2H  ATTCTAACATTGAATGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGCTACTAGATTTGACAAATTGTAAAGGTCTTCGTATAGATAATGGTGT 1360
RLG2I  ATTCTGGCATTGAATGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGCTACTGGATCTGACAGATTGTGGAGGTCTTCATATAGATAATGGCGT 1374
RLG2J  ATTCTAACATTGAATGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGCTACTAGATTTGACAAATTGTAAAGGTCTTCGTATAGATAATGGTGT 1358
RLG2K  ATTCTGGTATTGAGTGGTTGCCTTCCACAATCGGAAATTTGAAGGAGCTAAGGGTACTAGATTTGACAAATTGTGATGGTCTTCGTATAGATAATGGTGT 1374
RLG2L  ATTCTGGCATTGAACGGATACCTTCAGCAATCGGAAATTTGAAGAAGCTTAGGCAACTTGATCTGAGAGGTCGTTATGGTCTTTGTATAGAACAGGGTGT 1314
RLG2M  ATTCTCNCATTGAATGGTNGCCTTCCCCNGTCGGAAATTTAAAAAAGCTAAGGTTGCTTGATCTGAAATTTTGTGATNGTCTCCCTNTAGAACAAGGTGT 1349
```

```
              1410      1420      1430      1440      1450      1460      1470      1480      1490      1500
RLG2A   CTTAAAAAAATTGGTCAAACTGGAGGAGCTCTATAT-GACAGTG-------GTT----GAT-----------CGAGGTCGAAAGG--------CGA----  1437
RLG2B   CTTGAAAAGTTTTGTCAAACTTGAAGAATTTTATATTGGAGATGCATCTGGGTTTATAGATGATAACTGCAATGAGATGGCAGAGCGTTCTTACAACCTT  1475
RLG2C   CTTGAAAAGTTTGGTCAAACTTGAAGAATTTTATATTGGAAATGCATATG                                                   1417
RLG2D   CTTGAAAAGTTTGGTCAAACTTGAAGAATTTTATATTGGAAATGCATATGGGTTTATAGATG----------TAA------------------------  1411
RLG2E   CTTGAAAAATTTGGTGAAAATTGAAGAACTTTATATTGGTAGAGCAGATAT-TTTATAGAT-------------AGA-AGTA-A                 1432
RLG2F   CTTGAAAAATTTGGTGAAACTTGAAGAGCTTTATATTAGGAATG-------GTCTAC------------CAGT----TTACAGAGGAT------AA----  1437
RLG2G   CTTAAAAAAATTTGGTCAAACTTGAAGAGCTTTATATGGGTGTTA-ATC---GTCCGTATGGACAGGCCGT--TAGCTTGACAGATGAAAA          1466
RLG2H   CTTAAAAAAATTTGGTCAAACTTGAAGAGCTTTATATGGGTGTTA-ATC---ATCCGTATGGACT                                    1420
RLG2I   CTTAAAAAAATTTGGTCAAACTTGAAGAGCTTTATATGGGTGCTA-ATC---GTCTGTTTGGAAAG--TGC--CA--TT                      1442
RLG2J   CTTAAAAAAATTTGGTCAAACTTGAAGAGCTTTATATGGGTGTTA-ATC---GTCCGTATGGACAGGCCGT--TAGCTTGACAGATGAAAACT---GGCNA  1449
RLG2K   CCTAAAGAAATTGGTGAAACTTGAAGAGCTTTATAT---------------------GA-------------GAGTTGGTGG-----------------  1423
RLG2L   CTTGAAAAATTTGGTCGAACTTGAAGAACTTTATATTGGAAATGCATCTGCGTTTAGAGATTATAACTGCAATGAGATGGCAGA                 1398
RLG2M   CTTGAAAACTTTTGTCCA-CNTGAANAATTT-ATNTTGGAGACCC-TCTGGGTTTATNNATGAAAACTGCCATGANATGGCAGA----CCTTTCAACC--  1441

        XXXXXXXXXXXXXXXXXXXXXXXXX
                 1510      1520
RLG2A   -----------------------TTA - SEQ ID NO:27                                                           1439
RLG2B   TCTGCATTAGAATTCGCGTTCTTTA - SEQ ID NO:28                                                            1500
RLG2C                             - SEQ ID NO:29                                                            1417
RLG2D   -----------------------TG - SEQ ID NO:30                                                            1412
RLG2E                             - SEQ ID NO:31                                                            1432
RLG2F   -----------------------TA - SEQ ID NO:32                                                            1438
RLG2G - SEQ ID NO:33                                                                                        1466
RLG2H - SEQ ID NO:34                                                                                        1420
RLG2I - SEQ ID NO:35                                                                                        1442
RLG2J   TGGAAATGGCAGANCCTTTTTCACA - SEQ ID NO:36                                                            1474
RLG2K   TCGATATCAAAAGGC           - SEQ ID NO:37                                                            1437
RLG2L                             - SEQ ID NO:58                                                            1398
RLG2M   ------------------------A - SEQ ID NO:39                                                            1441
```

SEQ ID NO:

```
                 GETT----LKEVVEEKKMFNYIVEAVIGEKTDPIAIQQAVADYLGIELKESTKPARADKLREWFKAESDGGKNKFLVILDDVWQSVDLEDIGLSPFPNQG →40
                          10        20        30        40        50        60        70        80        90       100
RLG2A protein GKTTMMHRLKKVVKEKKMFNFIIEAVVGEKTDPIAIQSAVADYIGIEINEKTKPARTEKIRKWFVDNSGG--KKIIVIIDDVWQFVDINDIGLSPLPNQG 98 - 41
RLG2B protein GKTRMMQRLKKAXEEKKLFNYIVGAVIXEKTDPFAIQEAIADYIGIQINEKTKPARADKIREWFKKNSDGGKTKFLIVIDDVWQLVDLEDIGLSPFPNQG 100 - 42
RLG2C protein NTRX--AKAEEVAKKKEEFGYIIEAVIGEISDPIAIQQVVADYLCIELKESDKKTRAEKLRQGFKAKSDGGNTKFLIIILDDVWQSVDLXDIGLSPSPNQX 98 - 43
RLG2D protein EVAK--XK------RK--FGYIIEAVIXEISDPIAIQQVVADYLCIELKESDXKTRAEKLRQGFKAKSXGGNTKFLIIILDDVWQSVDLEXIGLSPSPNQG 90 - 44
RLG2E protein GRHND-AKVEEVAKENRMFNYMVEAVIGEKTDPLAIQQAVADYLCIELKESTKPARADKLREWFKANSGEGKNKFLVIFDDVWQSVDLEDIGLSHFPNQG 99 - 45
RLG2F protein LEDTMMQRLKKVVHEKKMFNFIVEAVIGEKTDPVAIQDAIADYLGVELNEKSKQARADKLRQGFKDKSDGGKNKFFVILDDVWQSVDLEDIGLSPFPNQG 100 - 46
RLG2G protein GRHDD-EELKEVVGQKKSFNIIIQVVIGEKTNPIAIQQAVADYLSIELKENTKEARADKLRKRF--EADGGKNKFLVIILDDVWQFVDLEDIGLSPLPNKG 97 - 47
RLG2H protein ---------KEVVERKKMFSIIVQVVIGEKINPIAIQQAVADYLSIELKENTKEARADKLRKWF--EADGGKNKFLVIIDDVWQFVDLEDIGLSPLPNKG 89 - 48
RLG2I protein CKKS----WKEVVEQKKTFNIIVQVVIGEKTNPIAIQQAVADSLSIELKENTKEARADKLRKWF--EADGGXNKFLVILDDVWXFVDLEXIGLSPHPNXG 94 - 49
RLG2J protein ERGR---------GXKKTFNIIVQVVIGEKTNPIAIQQAVADYLSIELKENTKEARADKLRKRF--EADGGKNKFLVILDDVWQFFDLEXIGLSPLPNKG 89 - 50
RLG2K protein LEDTMMXRLKNIIKEKRTFHYIVLVVIKENMDLISIQDAVADYLDMKLTESNESERADKLREGFQAKSDGGKNRFLIIILDDVWQSVNMEDIGLSPFPNQG 100 - 51
RLG2L protein ------------------FSYMVEAVIGEKTDPIAIQQAVXDYLRIQFKESTKPARADKLREWFKAXSXXGKNKFLVIFDDVWQSVDLEDIGLSPFPNQX 82 - 52
RLG2M protein AEE--------AAEEKKLFNYIVGAVIGEKTDPFAIQEAIADYLGIQLNEKTKPARADKLREWFKKNSDGGKTKFLIXLDDVWQLVDLEDIGLSPFPNQG 92 - 53

              VDFKVLLTSRDSHVCTVMGVEANSILNVGLLIEAEAQSLFQQFVETS--E---PELHKIGEDIVRKCCGLPIAIKTMACTLRNKRKDAWKDALSRLEHHD
                      110       120       130       140       150       160       170       180       190       200
RLG2A protein VDFKVLLTSRDKDVCTEMGAEVNSTFNVKMLIETEAQSLFHQFIEIS--DDVDPEIHNIGVNIVRKCGGLPIAIKTMACTLRGKSKDAWKNALLRLEHYD 196
RLG2B protein VDFKVLLTSRDSQVCTMMGVEANSIINVGLLTEAEAQSLFQQFVETS--E---PELQKIGEDIVRKCCGLPIAIKTMACXLRNKRKDAWKDALSRIEHYD 195
RLG2C protein VDFKVLLTSRDEHVCSVMGVEANSIINVGLLIEAEAQRLFQQFVETS--E---PELHKIGEDIVRRCCGLPIAIKTMACTLRNKRKDAWKDALSRLQHHD 193
RLG2D protein VXFKVLLTSRDEHVCSVMGVEANSIINVGLLIEXXXXRLFQQFVXTS--E---PELHKIXEXIVRRCCGLPIAIKTMACTLRNKRKDAWKDALSRLQHHD 185
RLG2E protein VDFKVLLTSR-DHVCXVMGVEANSILNVGLLVEAEAQSLFQQFVETF--E---PELHKIGEDIVRKCCGLPIAIKTMACTLRNKRKDAWKDALLHLEYHD 193
RLG2F protein VDFKVLLTSRDRHVCTVMGVEAKLILNVGLLIEAEAQSLFHQFVVTS--E---PELHKIGEDIVKKCFGLPIAIKTMACTLRHKRKDAWKDALSRLEHHD 195
RLG2G protein VNFKVLLTSRDSHVCTLMGAEANSIINIKVLKDVEGQSLFRQFAKNAGDDDLDPAFNGIADSIASRCQGLPIAIKTIALSLKGRSKSAWDVALSRLENHK 197
RLG2H protein VNFKVLLTSRDSHVCTLMGAEANSILNIKVLTAVEGQSLFRQFAKNAGDDDLDPAFNRIADSIASRCQGLPIAIKTIALSLKGRSKPAWDHALSRLENHK 189
RLG2I protein VXFKVLLTSRDSHVCTLMGAEANSILNIKVLKDVEGKSLFRQFAKNAGDDDLDPAFIGIADSIASRCQGLPIAIKTIALSLKGRSKSAWDVALSRLENHK 194
RLG2J protein VXFKVLLTSRXSHVCTLMGAEANSILNXKVLXDVQGKSLSAXLLKMRVXMTWIXLSLGXXIVLPVDVXGLXIAXKTIXXSLKGRSKSAWDVALSRLENHK 189
RLG2K protein VDFKVLLTSENKDVCAKMGVEANLIFDVKFLTEEEAQSLFYQFVKVS--DT---HLDKIGKAIVRNCGGLPIAIKTIANTLKNXNKDVWKDALSRIEHHD 195
RLG2L protein VDFKVLLTSRDXHVCTMMGVXANSXIXVGLLTEVEAXSLFXQFVETX--E---PELCKIXXVIVRKCCGLPIAIKTMACSLRNKRKDAWKDALSRIEHYD 177
RLG2M protein VDFKVLLTSRDSQVCTMMGVEANSIINVGLLXEAEAQSLFQQFVETS--E---PELQKIGEDIVRKCCGLPIAIKTMACTLRNKRKDAWKDALSRIEHYD 187
```

```
                 IGS--VAPKVFETSYDNLQDEETKSIFLMCGLFPEDFDIPTEELMRYGWGLKLFDRVYTIREARNRLNTCIERLVQTNLLIESDDVGCVKMHDLVRAFVL
                         210       220       230       240       250       260       270       280       290       300
RLG2A protein IFN--IVNGVFKMSYINLQDEEIKSTFILCGMYPEDFDIITEELVRYGWGLKLFKKVYTIGEARTRLNTCIERLIIFNLIMEVDDVRCIKMHDLVRAFVL 294
RLG2B protein IHN--VAPKVFETSYINLQEEETKSTFIMCGLFPEDFDIPTEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQTNLLIESDDVGCVKMHDLVRAFVL 293
RLG2C protein IGN--VATAVFRTSYENLPDKETKSVFLMCGLFPEDFNIPTEELMRYGWGLKLFDRVYTIIEARNRLNTCIDRLVQTNLLIGSDNGVHVKMHDLVRAFVL 291
RLG2D protein IGN--VATAVFXTSYENLPDKETKSVFIMCGLFPEDFNIPTEELMXYGWGLKLFDRVYTIIEARNRLNTCIERLVQANLLIGSDNGVHVKMHDLVRAFVL 283
RLG2E protein ISS--VAPKVFETSYHNLHNKETKSVFIMCGFFPEDFNIPIEELMRYGWGLKIFDRVYTIRQARIRINTCIERLVQTNLLIESDDGVHVKMHDLVRAFVL 291
RLG2F protein IQS--VVPKVFETSYNNLKDKETKSVFLMCGLFPEDLDIPIEELMRYGWGLRLFDRVNTITQARNRLNTCIERLVHTNLLIESVDGVHVKMHDLVRAFVL 293
RLG2G protein IGSEEVVREVFKISYDNLQDEVTKSIFLLCALFPEDFDIPTEELVRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVKMHDVVRDFVL 297
RLG2H protein IGSEEVVREVFKISYDNLQDEITKSIFLLCALFPEDFDIPTEELMRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVKMHDVVRDFVL 289
RLG2I protein IGSEEVVREVFKISYDNLQDEVTKSIFLLCALFPEDFDIPTEELVRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVKMHDVVRDFVL 294
RLG2J protein IGSEEVVREVFKXSYDNLQDEVTKSIFLLCALFPEDFDIPIEELVRYGWGLKLFIEAKTIREARNRLNNCTERLRETNLLFGSHDFGCVKMHDVVRDFVL 289
RLG2K protein IET--IAHVVFQMSYDNLQNEEAQSIFLLCGLFPEDFDIPTEELVXYGWGLRVFNGVYTIGEARHRINAYIELLKDSNLLIESDDVHCIKMHDLVRAFVL 293
RLG2L protein IXX--VAPKVFETKNINLHNKETKSAFLMCGLFPEDFNIPTEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQTNLLIESDDVGCVKMHDLVRAFVL 275
RLG2M protein IHN--VAPKVFETSYINLQEEETKSTFIMCGLFPEDFDIPTEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQTNLLIESDDVGCVKMHDLVRAFVL 285

              GMFSEVEHASIVNHGN--MPGWPEEND-IVHSCKRISLTCKGMSEFPVDLKFPNLTILKLMHGDKSLRFPQDFYEGMEKLQVISYDKMKYPLLPSSPQCS
                         310       320       330       340       350       360       370       380       390       400
RLG2A protein DMYSKVEHASIVNHSN--TLEWHADN--MHDSCKRLSLTCKGMSKFPTDLKFPNLSILKLMHEDISLRFPKNFYEEMEKLEVISYDKMKYPLLPSSPQCS 390
RLG2B protein GMFSEVEHASIVNHGN--MPGWPDENDMIVHSCKRISLTCKGMIEIPVDLKFPKLTILKIMHGDKSLRFPQDFYEGMEKLHVISYDKMKYPLLPLAPRCS 391
RLG2C protein GMYSEVEQASIVNHGN--MPGWPDENDMIVHSCKRISLTCKGMIEFPVDLKFPKLTILKIMHGDKSLKFPQEFYEGMEKLRVISYHKMKYPLLPLAPQCS 389
RLG2D protein GMYSEVEQASIVNHGN--MPGWPDENDMIVHSCKRISLTCKGMIEIPVDLKFPKLTILKLMHGDKSLKFPQEFYEGMEKLQVISYDKMKYPLLPLAPQCS 381
RLG2E protein VMFSEVEHASIINHGN--MLGWPENY--MTNSCKTISLTCKSMSEFPGDLKFPNLTILKLMHGDKLLRYPQDFYEGMEKLWVISYDEMKYPLLPSLPQCS 387
RLG2F protein GMFSEVEHASIVNHGN--MPEWTEND--MTDSCKQISLTCKSMLEFPGDLKFPNLKILKIMHGGKSLRYPQDFYQGMEKLEVISYDEMKYPLLPSLPQCS 389
RLG2G protein HIFSEVQHASIVNHGN--VSEWLEENH-SIYSCKRISLTCKGMSQFPKDLKFPNLSILKLMHGDKSLSFPENFYGKMEKVQVISYDKLMYPLLPSSLECS 394
RLG2H protein HIFSEVQHASIVNHGN--VSEWLEENH-SIYSCKRISLTCKGMSEFPKDLKFPNLSILKLMHGDKSLSFPENFYGKMEKVQVISYDKLMYPLLPSSLECS 386
RLG2I protein HIFSEVQHASIVNHGN--VSEWLEENH-SIYSCKRISLTCKGMSEFPKDLKFPNLSILKLMHGDKSLSFPENFYGKMEKVQVISYDKLMYPLLPSSLECS 391
RLG2J protein HMFSEVKHASIVNHGN--MSEWPEKND-TSNSCKRISLTCKGMSKFPKDINYPNLLILKLMHGDKSLCFPENFYGKMEKVQVISYDKLMYPLLPSSLECS 386
RLG2K protein DTFNRFKHSLIVNHGNGGMLGWPEN-DMSASSCKRISLICKGMSDFPRDVKFPNLLILKLMHADKSLKFPQDFYGEMKKLQVISYDHMKYPLLPTSPQCS 392
RLG2L protein GMYSEVEHASIVNHGN--MHGWTKND--MNDSCKTVSLTCESVSEFPGDLKFPNLKLLKLMHGDKMLRFSQDFYEGMEKLQVISYHKMKYPLLPSSPQCS 371
RLG2M protein GMFSEVEHASIVNHGN--MPGWPDENDMIVHSCKRISLTCKGMIEIPVDLKFPKLTILKLMHGDKSLRXPXDFYEGMEKLXVISYDKMKXPLLPLAPRCS 383
```

```
                 TNLRVLHLHECSLRMFDCSSIGNLLNLEVLSFANSGIEWLPSTIGNLKKLRLLDLTNCYGLRIENGVLKNLVKLEELYIGNANG-FG-------------
                         410       420       430       440       450       460       470       480     . 490
RLG2A protein VNLRVFHLHKCSLVMFDCSCIGNLSNLEVLSFADSAIDRLPSTIGKLKKLRLLDLTNCYGVRIINGVLKKLVKLEELYMIVV----------DRGRKAI  479
RLG2B protein TNIRVLHLTECSLKHFDCSSIGNLSNLEVLSFANSHIEWLPSTVRNLKKLRLLDLRFCDGLRIEQGVLKSFVKLEEFYIGDASG-FIDDNC  481
RLG2C protein TNIRVLHLTECSLKMFDCSCIGNLSNLEVLSFANSCIEWLPSTVRNLKKLRLLDLRLCYGLRIEQGVLKSLVKLEEFYIGNAY  472
RLG2D protein TNIRVLHLTECSLKMFDCSSIGNLSNLEVLSFANSRIEWLPSTVRNLKKLRLLDLRFCDGLRIEQGVLKSLVKLEEFYIGNAYG-FID----VM  470
RLG2E protein INLRVLHLHRCSLMMFDCSCIGNMLNLEVLSFVKSGIEWLPSTIGNLKKLRLLDLRDCYGLRIEKGVLKNLVKIEELYIGRA-DI  471
RLG2F protein TILRVLHLHECSLRMFDCSSIGNLFNMEVLSFANSSIELLPSVIGNLKKLRLLDLTNCYGVRIEKDVLKNLVKLEELYIRNGLPVYRG  477
RLG2G protein TNVRVLHLHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKLRLLDLTNCKGLRIDNGVLKNLVKLEELYMG-VNRPYG----QAVSLTDE  488
RLG2H protein TNVRVLHLHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKLRLLDLTNCKGLRIDNGVLKNLVKLEELYMG-VNHPYG  472
RLG2I protein TNLRVLHLHECSLRMFDCSSIGNLLNMEVLSFANSGIEWLPSTIGNLKKLRLLDLTDCGGLHIDNGVLKNLVKLEELYMG-ANRLFG----KC-----H  480
RLG2J protein TNVRVLHLHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKLRLLDLTNCKGLRIDNGVLKNLVKLEELYMG-VNRPYG----QAVSLTD  479
RLG2K protein TNLRVLHLHQCSL-MFDCSSIGNLLNLEVLSFANSGIEWLPSTIGNLKELRVLDLTNCDGLRIDNGVLKKLVKLEELYMRVG------------GRYQK  478
RLG2L protein TNLRVLHLHRCSLRMLDCSCIGNLTNLEVLSFANSGIERIPSAIGNLKKLRQLDLRGRYGLCIEQGVLKNLVELEELYIGNA-SAFRDYNCNEMA  465
RLG2M protein TNIRVLHLTECSLKMFDCSSIGNLSNLEVLXFANSXIEWXPSPVGNLKKLRLLDLKFCDXLPXEQGVLKTFVHXXNLXW-RPSG-FXXENCHXMADLST  480
```

```
                                                                                                                  SEQ ID NO:
           TGTGAGACCGTGACTTGGATGGTAGATAAATTTAGTAAACTTAACC-CTTC--AATTAACCTACCTTTTTCTTATTAACTCAATTTCAAGCTAAATTCTG  → 54
                    10        20        30        40        50        60        70        80        90       100
AC15-2A    TGTGAGACCGTGACTTGGATGGTAGATAAATTTAGTAAACTTAACC-CTTC--AATTAACCTACCTTTTTCTTATTAACTCAATTTCAACCTAAATTCTG 97 - 56
AC15-2B    TGKGAGACCGTGACTTGGATGGTAGATAAATTTAGTAAACTTAACC-CTTC--AATTAACCTACCTTTTTCTTATTAACTCAATTTCAAGCTAAATTCTG 97 - 57
AC15-2C    -GTGAGACCGTGACTTGGATGGTAGATAAATTTAGTAAACTTAACC-CTTC--AATTAACCTACCTTTTTCTTATTAACTCAATTTCAACCTAAATTCTG 96 - 58
AC15-2D    -GTGAGACCGTGACTTGGATGGCAGATAAATTTAGTAAACTTAACC-CTTC--AATTAACCTACCTTTTTCTTATTAACTCAATTTCAAGCTAAATTCTG 96 - 59
AC15-2E    TGTGAGACCGTGACTTGGATG------------------------------------------------------TGAAGATAAATTTA-G-TAAACTCCA 45 - 60
AC15-2G    TGTGAGACCGTGACTTGGATGGCAGATAAATTTAGTAAACTTAACC-CTTC--AATTAACCTACCTTTTTCTTATTAACTCAATTTCAACCTAAATTCGG 97 - 61
AC15-2H    -GTGAGACCGTGACTTGGATGGTAGATAAATTTAGTAAACTTAACC-CTTC--AATTAACCTACCTTTTTCTTATTAACTCGATTTCAAGCTAAATTCTG 96 - 62
AC15-2I    -----------------------AGATAAATTTAGTAAACTTAACC-CTTC--AATTAACCTACCTTTTTCTTATTAACTCAATTTCAAGCTAAATTCTG 74 - 63
AC15-2J    C-CAAAACACAAATTGAAAACCNGGATCATCCAAATAACATCATCCACTTCCAAATGAGCCCAAATTCAATTATTCAA--GGNTTCTAAGCCTGTTAATG 97 - 64
AC15-2L    ---GA-----TAATTTAG--------------------------------------------------------TAAACTCAATTTCAAG-CCAACTCTG 35 - 65
AC15-2N    TGTGAGACCGTGACTTGGATGGTAGATAAATTTAGTAAACTTAACC-CTTC--AATTAACCTACCTTTTTCTTATTAACTCAATTTCAAGCTAAATTCTG 97 - 66
AC15-2O    AG-AGCAGAGCAGTATGGATTTCATTTCACTTT------CTACTTACTTAAGGATTAGCTTCTGTTTTTTTGAATAA---AA----AAGGGGACATCT- 85 - 67

           ATTCTTGTTTGAAAATAAGTT-GCATCTTTATTTT-TG--TATTATCTTGTTGCATAGGATCCT-TAGCATCTTTTAATAATTTATTT-----GAAGGTG
                   110       120       130       140       150       160       170       180       190       200
AC15-2A    ATTCTTGTTTGAAAGTAAGTT-GCATCTTTATTTT-TG--TATTATCTTGTTGCATAGGATCCT-TAGCATCTTTTAATAATTTATTT-----GAAGGTG 187
AC15-2B    ATTCTTGTTKGAAAGTAAGTT-GCATCTTTATGTT-TG--TATTATCTTGTTGCATAGGATCCT-TAGCATCTTTTAATAATTTATTT-----GAAGGTG 187
AC15-2C    ATTCTTGTTTGAAAATAAGTT-GCATCTTTATTTT-TG--TATTATCTTGTTGCATAGGATCCT-TAGCATCTTTTAATAATTTATTT-----GAAGGTG 186
AC15-2D    ATTCTTGTTTGAAAATAAGTT-GCATCTTTATTTT-TGCATATTATCTTGTTGCATAGGATCCT-TAGCATCTTTTAATAGTTTATTT-----GAAGCTG 188
AC15-2E    TCTCTTGTTTGAAAAT-AATCCACATCTT-ACTTAAAATATATTTAT-TGTTGCATAGAAGCCTATAACATCTTCTAATAATCCCTCT-----GAAGGTG 137
AC15-2G    ACTCTTGTTTGAAAATAAGTT-GCATCTTTATTTT-TG--TATTATCTTGTTGCATAGGATCCT-TAGCATCTTTTAATAATTTATTT-----GAAGGTG 187
AC15-2H    ATTCTTGTTTGAAAATAAGTTAGCATATTTATTTT-T-TGTATTATCTTGTTGCATAGGATCCT-TAGCATCTTTTAATAAT----------------- 176
AC15-2I    ATTCTTGTTTGAAAATAAGTT-GCATTTTTATGTT-TG--TATTATCCTGTTGCATAGGATCCT-TAGCATCTTTTAATAATTTATTT-----GAAGGTG 164
AC15-2J    CCCATTCTTGGATTACCACCCCGCAATGGGAAACGATTCAAAACAGGGCGTTACATA--ATTTG-TTGTGGTTTTGTATTTTTTATTTCCGGTGAAGGTG 194
AC15-2L    TCTCTTGTTTGAAAAT-AGTTCACATCTT-AATTAAAATATGTTTTCATGTTTCATAGGAGCATATAACAGCTTTCAATTATTTATCT----------- 122
AC15-2N    ATTCTTGTTTGAAAGTAAGTT-GCATCTTTATGTT-TG--TATTATCTTGTTGCATAGGATCCT-TAGCATCTTTTAATAATTTATTT-----GAAGGTG 187
AC15-2O    --TCT--------AATAA-TGCACATCTTAAATTAAAAAGTATTTAATTGTTGCATAGCAGCGTATAACATCTTCTAATAATTTATCT-----GAAGGTG 168
```

```
                      210       220       230       240       250       260       270       280       290       300
AC15-2A    AAAGATCCAACTATTTTTAATCTGTTGGCATTTTCCATCATTTGCAACTGTTTCTTGAAAAAAAAA-TACCTAAAATCAAAATAACCATTTTCAAATCCA 286
AC15-2B    AAAGATCCAACTATTTTTAATCTGTTGGCATTTTCCATCATTTGCAACTGTTTCTTGAAAAA---AATACCTAAAATCAAAATAACCATTTTCATATCCA 284
AC15-2C    AAAGATCCAACTATTTTTAATCTGTTGACGTTTTCCATCATTTGCAACTGTTTCTTGAAAAAAAAA-TACCTAAAATCAAAATAACCATTTTCAAATCCA 285
AC15-2D    AAAGATCCAACTAGTTTTGATCTGTTGGCATTTTCCATCATTTGCAACTGTTTCTTGAAAAAAAA--TACCTAAAATCAAAATAACCATTTTCAAATCCA 286
AC15-2E    AGAGATCAAACTATTTGTAATCTGTTGACAATTTCAATCATTTGCAATTGTTCCTTGAAAAAT----TACTTAGAATCAAAGAAATATTTTTCTAATCCA 233
AC15-2G    AAAGATCAAACTATTTTTTAGCTGTTGGCATTTTCCATCATTTGCAACTGTTTCTTGAAAAAAAAAATACCTAAAATCAAAATAACCATTTTCAAATCCA 287
AC15-2H    ---------------------CTGTTGGCATTTTCCATCATTTGCAACTGTTTCTTGAAAAAAAAA-TACCTAAAATCAAAATAACCATTTTCAAATCCA 253
AC15-2I    AAAGATCCAACTATTTTTTAGCTGTTGGCATTTTCCATCATTTGCAACTGTTTCTTGAAAAAAAAA-TACCTAAAATCAAAATAACCATTTTCAAATCCA 263
AC15-2J    AAAGATCCAACTATTTTTAATCTGTTGGCATTTTCCATCATTTGCAACTGTTTTTTGAAAAAAAAA-TACCTAAAATCAAAATAACCATTTTCAAATCCA 293
AC15-2L    ------CCAACACTTTTTAATTTGTTGAGAATTTTCATCATATGCAAATGTTTCTTGAAAAAT----TGCCTAAAATCAA---AATATTTTTCAAATCCA 208
AC15-2N    AAAGATCCAACTATTTTTAATCTGTTGGCATTTTCCATCATTTGCAACTGTTTCTTGAAAAA---AATACCTAAAATCAAAATAACCATTTTCATATCCA 284
AC15-2O    AAAGATCCAACTACTTCTAATTTGTTAACAATTTCAATCATTTGCAAATGTTCCTTAAAAAATTAATTACCTGAAATCAAAACAATCTTCTTCAAATCCA 268

           AAATTATAAGAGAGAATTGTTAATGGAC-ATGGAATCATAAATCATTAAC---ACAGTTCAGTACACAAGTTGCTAATTACATTTCTTGCTG--------
                      310       320       330       340       350       360       370       380       390       400
AC15-2A    AAATTATAAGAGAGAATTGTAAATGGAC-ATGGAATCATAAATCATTAAC---ACAGTTCAGTAAACAAGTTGCTAATTACATTTCTTGCTG-------- 375
AC15-2B    AAATTATAAGAGAGAATTGTTAACGGAC-ATGGAATCATAAATCATTAAC---ACAGTTCAGTACACAGGTTGCTAATTACATTTCTTGCTG-------- 373
AC15-2C    AAATTATAAGAGAGAATTGTAAATGGAC-ATGGAATCATAAATCATTAAC---ACAGTTCAGTAAACAAGTTGCTAATTACATTTCTTGCTG-------- 374
AC15-2D    AAATTATAAGAGAGAATTGTTAATGGAC-GTGGAATCATAAATCATTAAC---ACAGTTCAGTACACAAGTTGCTAATTACATTTCTTGCTG-------- 375
AC15-2E    AAATTATAAGAGAGAATTGGGAATGGACAGTGTAATTATAAATCATTAAC---ACAATTCAATATTCAAGTTACTAATTACATTTATTGTTGGGATATAT 330
AC15-2G    AAATTATAAGATAGAATTGTTAATGGAC-ATGGAATCATAGATCATTAAC---ACAGTTCAGTAAACAAGTTGCTAATTACATTTCTTGCTG-------- 376
AC15-2H    AAATTATAAGAGAGAATTGTTAATGGAC-ATGGAATCATAAATCATTAAC---ACAGTTCAGTACACAAGTTGCTAATTACATTTCTTGCTG-------- 342
AC15-2I    AAATTATAAGAGAGAATTGTAAATGGAC-ATGGAATCATAAATCATTAAC---ACAGTTCAGTAAACAAGTTGCTAATTACATTTCTTGCTG-------- 352
AC15-2J    AAATTATAAGAGAGAATTGTAAATGGAC-ATGGAATCTTAAATCATTAAC---ACAGTTCAGTACACAAGTTGCTAATTACATTTCTTGCTG-------- 382
AC15-2L    AAATTCTAAGAGATAATTGTGAATGGACT-TGAAGTTATAAATCATTTACTACACACTTCATTAGTGAAGTTACTAATTACATATCTTGTTGGGATATAT 307
AC15-2N    AAATTATAAGAGAGAATTGTTAACGGAC-ATGGAATCATAAATCATTAAC---ACAGTTCAGTACACAGGTTGCTAATTACATTTCTTGCTG-------- 373
AC15-2O    AAATTATGAGACAGAATTGAGAAGGGAT-GTGAAATTATAAACCATTAAC---ACAATTCCATGCTCACGTTACTAATTACATTTCTTGTTGGGATATAT 364
```

```
               ---TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGCCACTGGCA---GTATTTCTAAT---GTTGTATT---CCCATCCTGTCTCATGCA
                       410       420       430       440       450       460       470       480       490       500
AC15-2A        ---TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGCCACTGACA---GTATTTCTAAT---GTTGTATT---CCCATCCTGTCTCATGCA 462
AC15-2B        ---TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGCCACTGGCA---GTATTTCAAAT---ATTGTATT---CCCATCCTGTCTCATGCA 460
AC15-2C        ---TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGCCACTGGCA---GTATTTCAAATC---TTGTATT---CCCATCCTGTCTCATGCA 461
AC15-2D        ---TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGTCACTGATACTAATATTTCTAATGATGTTGTATTATTCCCATCCTGTCTCATGCA 471
AC15-2E        GTGTGCAGACTGATACTCTGTCAGAGGAAGAGATATTACAAGAAGTCACTGGTA---GTATTTCTAA---TGTTGCAGT---CCCATCCAGTCTCATACA 421
AC15-2G        ---TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGCCACTGACA---GTATTTCTAAT---GTTGTATT---CCCATCCTGTCTCATGCA 463
AC15-2H        ---TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGCCACTGGCA---GTATTTCTAAT---GTTGTA---TTCCCACCCTGTCTCATGCA 429
AC15-2I        ---TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGCCACTGGCA---GTATTTCAAAT---CTTGTATT---CCCATCCTGTCTCATGCA 439
AC15-2J        ---TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGCCACTGACA---GTATTTCTAA---TGTTGTATT---CCCATCCTGTCTCATGCA 469
AC15-2L        ATGCACAGACTGATATTCTGTCAGAGGAAGAGACATTGCGAGAAATCACTGGCA---ATATTTCTAA---TGTTGCATT---CCCATCCTGTCTCATACC 398
AC15-2N        ---TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGCCACTGGCA---GTATTTCAAAT---ATTGTATT---CCCATCCTGTCTCATGCA 460
AC15-2O        ATGTACAGACTGATATTTTGTCAGAGGAAGTGAAATTACAAGAAGTCACTGATACTA---TTTCTA---ATGTTG---TATTCACATCGTGTCTCATACA 455

               CTCTTTTCATAAC---CTCCATAAACTTAAATTGAAGAGATTTGAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTGAGAGTCCAACAAGTAGAGAATTG
                       510       520       530       540       550       560       570       580       590       600
AC15-2A        CTCTTTTCATAAC---CTCCAGAAACTTATATTGAACAGAGTTAAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTGAGAGTCCAACAAGTAGAGAATTG 559
AC15-2B        CTCTTTTCATAAC---CTCCATAAACTTAACTTGAACAGAGTTGAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTGAGAGTCCAACAAGTAGAGAATTG 557
AC15-2C        CTCTTTTCATAAC---CTCCGTGTGCTTACATTGGATAATTATGAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTGAGAGTCCAACAAGTAGAGAATTG 558
AC15-2D        CTCTTTTCATAAC---CTCCATAAACTTAAATTGGAAAATTATGAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTGAGAGTCCAACATGTAGAGAATTG 568
AC15-2E        CTCTTTTCCTAAC---CTCCGTAGACTTGAATTGGAGAAATATAAGGGAGTGGAGGTTGTGTTTGAAATAGAGAGT------CCCACAAGTAGAGAATTG 512
AC15-2G        CTCTTTTCATAAC---CTCCATAAACTTAAATTGAAGAGAGTTAAAGGAGTGGAGGTGGTGTTTGAGATAGAGGGTGAGAGTCCAACAAGTAGAGAATTG 560
AC15-2H        CTCTTTTCATAAC---CTCCATAAACTAAAAATGAAGAAGTATAAAGGAGTGGAGGTGGTGTTTGAGATAGAG------AGTCCAACAAGTAGAGAATTG 520
AC15-2I        CTCTTTTCATAAC---CTCCATAAACTTAACTTGAACAGAGTTGAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTGAGAGTCCAACAAGTAGAGAATTG 536
AC15-2J        CTCTTTTCATAAC---CTCCAGAAACTTATATTGAACAGAGTTAAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTGAGAGTCCAACAAGTAGAGAATTG 566
AC15-2L        CTCTTTTCATAAC---CTCCATAAACTTTACTTGAAGAAGTATGAAGGAGTGAAGGTGGTGTTTGAGATGGAGAAT------CCAACAAGTAGAGAATTG 489
AC15-2N        CTCTTTTCATAAC---CTCCATAAACTTAACTTGAACAGAGTTGAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTGAGAGTCCAACAAGTAGAGAATTG 557
AC15-2O        CTCTTTTTATAACAACCTCCGTAAACTCAACTTGGAGAAGTATGGAGGAGTTGAGGTTGTGTTTGAGATAG------AGAGTTCAACAAGTAGAGAATTG 549
```

```
                 610       620       630       640       650       660       670       680       690       700
AC15-2A   GTAACAACTCACCATAACCAACAACAAC--CTAT----TATACTTCCCAACCTCCAGGAATTGATTCTATGGAATATGGACAACATGAGTCATGTGTGGA 653
AC15-2B   GTAACAACTCGCAATAACCAACAACAGC--CTAT----TATACTTCCCTACCTCCAGGATTTGTATCTAAGGAATATGGACAACACGAGTCATGTGTGGA 651
AC15-2C   GTAACAACTCACAATAACCAACAACAGC--CTAT----TATACTTCCCTACCTCCAGGAATTGTATCTAAGGAATATGGACAACACGAGTCATGTGTGGA 652
AC15-2D   GTAACAACTCACAATAACCAACAACAGC--CTAT----TATACTTCCCAACCTCCAGGAATTGTATCTAAGGAATATGGACAACACGAGTCATGTGTGGA 662
AC15-2E   GTAACAACTCAACATAGTCAACAACCAC-----T-------ACTTCTCAACCTTGAGGAATTGCATCTAAGTTTTATGGGAAGCATGAGTCATGTGTGGA 600
AC15-2G   GTAACAACTCACAATAACCAACAACAGC--CTAT----TATACTTCCCTACCTCCAGGAATTAGTTCTAAGGAATATGGACAACACGAGTCATGTGTGGA 654
AC15-2H   GTAACAACTCACCATAACGAACAACATC--CTAT----TATACTTCCTAACCTCCAGCATTTGGATCTAAGGAATATGGACAACATGATTCATGTGTGGA 614
AC15-2I   GTAACAACTCACCATAACCAACAACAGC--CTGT----TATATTTCCCAACCTCCAGCATTTGGATCTAAGGGGTATGGACAACATGATTCACGTGTGGA 630
AC15-2J   GTAACAACTCACCATAACCAACAACAGC--CTGT----TATATTTCCCAACCTCCAGCATTTGGATCTAAGGGGTATGGACAACATGATTCGCGTGTGGA 660
AC15-2L   GTAACAACTCACCATAACCAACAACAGC--CTAT-------ACTTCTCAACCTCCAGGAATTATATCTATATAATATGGACAACATGAGCCATGTATGGA 580
AC15-2N   GTAACAACTCACCATAACCAACAACAAC--CTAT----TATACTTCCCAACCTCCAGGAATTGATTCTATGGAATATGGACAACATGAGTCATGTGTGGA 651
AC15-2O   GTAACAACATACCATAAACAACAACAACAACAACCTATATTTCCCAACCTTGAGGAATTATATCTATATTATATGGACAACATGAGTCATGTATGGA 649

          AGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAATCCCCATTCCACAACCTCACAACCATAAATATTTATTAATGCAAAAGCAT
                 710       720       730       740       750       760       770       780       790       800
AC15-2A   AGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAATCCCCATTCCACAACCTCACAACCATAAAAATTATGTATTGCAAAAGCAT 753
AC15-2B   AGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAATCCCCATTCCACAACCTCACAACCATAAATATTCTTAAATGCAAAAGCAT 751
AC15-2C   AGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAATCACCATTCCACAACCTCACAACCATAGAAATGAGATGGTGTCATGGCTT 752
AC15-2D   AGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAATCACCATTCCACAACCTCACAACCATAGAAATGAGATGGTGTCATGGCTT 762
AC15-2E   AGTG---CAACTGGAATAAATTCTTCACTCTTTCAAAACAA---TCTGAATCCCCATTCCAGAACCTCACAGCCATATACCTGATGGACTGCAGAAGCAT 694
AC15-2G   AGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAATCCCCATTCCACAACCTCACAACCATAAATATTTACAGATGCAAAACCAT 754
AC15-2H   AGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAATCCCCATTCCACAACCTCAGTAACATACATATTTATGAATGCAAAAACAT 714
AC15-2I   AGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAATCCCCATTCCACAACCTCACAACCATAAATATTGAGTTTTGC          722
AC15-2J   AGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAATCCCCATTCCACAACCTCACAACCATAAATATTGATTTTTGCAGAAGCAT 760
AC15-2L   AGTG---CAACCGGTATAGGTTCTTCAATCTTCCAAAACAACAATCTGAATCAACATTCCACAACCTCACAACCATACATGTGTTCCAATGCAAAAGCAT 677
AC15-2N   AGTGCGGCAACTGGAATAAATTCTTCACTCTTCCAAAAGAACAATCAGAATCCCCATTCCACAACCTCAGTAACATACATATTTATGAATGCAAAAGCAT 751
AC15-2O   AGTGCAACAACTGGAATAAATTTTT------------ACAACAATCAGAATCCCCATTCCACAACCTCACAACCATACACATGTCCGATTGCAAAAGCAT 737
```

SEQ ID NO:

	810 820		
AC15-2A	TAAGTACTTGTTTTCACCTCTCACGG	-56	779
AC15-2B	TAAGTACTTGTTTTCACCTCTCACGG	-57	777
AC15-2C	TAGGTACTTGTTT-CACCTCTCACGG	-58	777
AC15-2D	TAGGTACTTGTTTTCACCTCTCACGG	-59	788
AC15-2E	TAAGTACTTGTTTTCACCTCTCACGGA	-60	721
AC15-2G	TAAGTACTTGTTTTCACCTCTCACGGA	-61	781
AC15-2H	TAAGTACTTGTTTTCACCTCTCMC	-62	738
AC15-2I		-63	722
AC15-2J	TAAGTACTTGTTTTCACCTCTCAC	-64	784
AC15-2L	TAAGTACTTGTTTTCACCCCTC	-65	699
AC15-2N	TAAGTACTTGTTTTCACCTCTCACGGA	-66	778
AC15-2O	TAAGTACTTGTTTTCACCTCTCACGG	-67	763

SEQ ID NO:68

RLG3 (real RLG3)

[Strand]

```
1     AATGGCAAAA GAAGTCGGAG CAAGAGCTAA GTTAGAGCAT CTATTTGACG TCATTATCAT GGTAGATGTC
71    ACTCAAGCAC CCAACAAGAA CACAATTCAA AGTAGTATTT CAGAACAGTT GGGATTAAAA CTGCAAGAAG
141   AGAGCTTGTT GGTAAGAGCA GCTAGGGTAA GTGCGAGGTT AAAAATGCTT ACAAGGGTGC TGGTGATATT
211   AGACGATATA TGGTCAAGGC TTGACATGGA GGAACTTGGG ATTCCCTTTG GATCAGATAG ACAACACCAC
281   GGCTGCAAAA TCTTGTTGAC TTCAAGAAGT ATTAGTGCTT GTAACCAGAT GAGAGCTGAT AGAATCTTTA
351   AAATACGAGA AATGCCACTG AATGAAGCAT GGCTTCTTTT CGAAAGAACA GCTAAAAAAG CTCCGAATCT
421   GCATCAAGTA GCAAGAGATA TCGTGGAGGA GTGTGGTGGG C
```

RLG4

SEQ ID NO:69

```
1    GAATTCGGTG TTGGTAAGAC AACTCTTGCC TCTTCTGTTT ATGATGAAAT CTCTAGCAAG TTTGATGGTT
71   GCTGCTTTCT AAAAATATCT GGGAGGAATC AAGTAATAAA GACGGTATAG AAAGATTGCA AGAAAAAATC
141  ATTTGTGATG TTTTGAAACA AGAGCAAGTG GGCGTAGGGA GAGTTGAAGA AGGAAAGCGC ATGATAAAGG
211  ATAGGTTACA ACATAGAAAG GTATTGATTG TGCTTGATGA TGTCGACAAC GTTGAGCAGC TAGCTAGAAC
281  AGTTGGCTGG ATCACATGAT TGGTTTGGTG AAGGTAGCCG CATAATAATC ACAACTAGAG ATGAACATGT
351  ATTAATTGCA CACAAAGTAG ATGTGATACA CAATATAAGC TTGTTAAACA ACGATGAAGC TATGCATCTC
421  TTCTGCAAGC AAGCACCACG GGGTCACAAA CGTATACAAG ATTATGAGCA ACTTTTAAAA CATGTGGTTT
491  CTTATGCTGG TGGGCTTCCA CTAGCACTGT CGAC
```

SEQ.ID NO: 70

RLG1-E169

[Strand]

1 ATCGTAACCG TTCGTACGAG ANCGCTGTCC CTCCTTCATC TTTTGTCATA TGTCATATTC TCATNNATTN
71 TGCCACATNT AATTTTGTGG TTATTTTAAA TTAATTTTTA TTCCACATGT CATTTTATGA GTTTTTCTAT
141 TTTATTGAGT TTCACATAAT ATTTAAATGT AATAACAATA AATGCATATT TATTTTTCTT TAAATAAACG
211 CATATAATAT ATAGATTAAA ATCATATAAT ACATAGGTTA AACTCATATA ATACATATGT TCATCCCCAG
281 TTTATTTATA TGTCTCATCC TTAATTTATT TATTATTTAT TTATTAGAGT AGATGATCTT TGTGATATTA
351 AAAATTTAAT TTGTTCAAAA TTTAAAATTA TTAATAATCC CACAATTTGA ATAAAATTAA AAAAAATGGN
421 CCCACCATTA GTCCATCACT TTTTCAGCTC ATCAATATCG TGAGTATTCT CCTTCGTTTC CACCCTAATC
491 AATATTTCCA GCGAATGACA GACTCCTACG GCGTTTCTGA ATTTGCGTTC CGACACTGTT CATTGAAGGA
561 GATAATAAAT CAAATGGAGC TGCTCCAATG TTCATTGCTG ATGAAAGGTG AATTGTATGT GAAGANAATG
631 TCAGCGATCN ATCTCCATCC GGAACCCACC ACATTATCAG TGTACCACCA AACCACTCAA AACGGYGGAA
701 GTAGRRAKAC WRKAAAGTCA TGAAGAATAG ATTATTTTTG TCCTCATGGG CTGACTGAGG AGCGGGTTTA
771 GTTCATCATT TTTCTTTGAN CAAAGAATTA TCGGTCCATC GAATTTTTAC ATCGACAAAG AAGTTTCACT
841 TCGCAATGTT TTGTTAAACA ATTTTTAATC TTTTTATCTT TTCGTTGAAA CTCCTCAATT GCAACTTGCA
911 ACTTGCAACT TTTGGGCCCA CAAATTTGTG GTGGGCGTTA ATTTAATCCA CATATTCACT GTAAACAATA
981 ATTCAAATCG ATCTCTGTTC ATCCAATTCA TCAACATCTC TTGATAATTG AAATCATTCA CGCTTCATCC
1051 ATTTCATCCA CATCTATACT ATATTCTCTG CTCTTATCAT ATTAAACGAT GGCTGAAATC GTTCTTTCTG
1121 CCTTCTTGAC AGTGGTGTTT GAAAAGCTGG CATYTGAAGC CTTGAAGAAG ATTGTTCGCT CCAAAAGAAT
1191 TGAATCTGAG CTTAAGAAAT TGAAGGAGAC ATTAGACCAA ATCCAAGATC TGCTTAACGA TGCTTCCCAG
1261 AAGGAAGTAA CTAATGAAGC CGTTAAAAGA TGGCTGAATG ATCTCCAACA TTTGGCTTAT GACATAGACG
1331 ACCTACTTGA TGATYTTGCA ACTGAAGCTG TTCAWCGTGA GTTGACCGAG GAGGGTGGAG CCTCCTCCAG
1401 TATGGTAAGA AAACTAATCC CAAGTTGTTG CACAAGTTTC TCACAAAGTA ATAGGATGCA TGCCAAGTTA
1471 GATGATATTG CCACCAGGTT ACAAGAACTG GTAGAGGCAA AAAATAATCT TGGTTTAAGT GTGATAACAT
1541 ATGAAAAGCC AAAAATTGAA AGGTATGAGG CGTCTTTGGT AGATGAAAGC GGTACTGTCG GACGTGAAGA
1611 TGATAAGAAA AAATTGCTGG AGAAGCTGTT GGGGGATAAA GATGAATCAG GGAGTCAAAA CTTCAGCATC
1681 GTGCCCATAG TTGGTATGGG TGGAGTTGGT AAAACAACTC TAGCTAGACT TTTGTATGAT GAAAAGAAAG
1751 TGAAGGATCA CTTCGAACTC AGGGCTTGGG TTTGTGTTTC TGATGAGTTC AGTGTTCCCA ATATAAGCAG
1821 AGTTATTTAT CAATCTGTGA CTGGGGAAAA GAAGGAGTTT GAAGACTTAA ATCTGCTTCA AGAAGCTCTT
1891 AAAGAGAAAC TTAGGAACCA GCTATTTCTA ATAGTTTTGG ATGATGTGTG GTCTGAAAGC TATGGTGATT
1961 GGGAGAAATT AGTGGGCCCA TTCCTTGCGG GGTCTCCTGG AAGTAGAATA ATCATGACAA CTCGGAAGGA
2031 GCAATTGCTC AGAAAGCTGG GCTTTTCTCA TCAAGACCCT CTGGAGGGTC TATCACAAGA TGATGCTTTG
2101 TCTTTGTTTG CTCAACACGC ATTTGGTGTA CCAAACTTTG ATTCACATCC AACACTAAGG CCACATGGAG
2171 AACTGTTTGT GAAGAAATGT GATGGCTTAC CTCTAGCYTT AAGAACACTT GGAAGGTTAT TAAGGACAAA
2241 AACAGACGAG GAACAATGGA AGGAGCTGTT GGATAGTGAG ATATGGAGGT TAGGAAAGAG CGATGAGATT
2311 GTTCCGGCTC TTAGACTAAG CTACAATGAT CTTTCTGCCW CTTTGAAGCT RTTRTTTGCA TAYTGCTCCT
2381 TGTTTCCCAA GGACTATGAG TTTGACAAGG AGGAGTTGAT TCTATTGTGG ATGGCAGAAG GGTTTTTGCA
2451 CCAACCAACT AYAAACAAGT CAAAGCAACG KTTGGGTCTT GAATATTTTR AAGAGTTRTT GTCAAGRTCR
2521 TTTTTTCAAC ATGCTCCTAA TRRCAAATCS TTGTTTGTGA TGCATGACCT AATGAATGAT TTGGCTACAT
2591 TTGTTGCTGG AGAATTTTTT TCAAGGTTAG ACATAGAGAT GAAGAAGGAA TTTAGGATGS AATCTTTGGA
2661 RAAGCACCGN CATATGTCAT TTGTATGTGA GRATTACATA GGTTACAAAA RGTTCGAGCC ATTTAGAGGA
2731 GCTAAAAATT TGAGAACATT TTTAGCATTG TCTGTTGGGG TGGTAGAAGA TTGGAAGATG TTTTACTTAT
2801 CAAACAAGGT CTTGAATGAC WTACTTCARG ATTTACCATT GTTAAGGGTC CTRAKTTTGA TTRRTCTTAY
2871 AATAASYRAG GTACCARAAK TCGTSGGTAG TATGAASCAC TTGCGGTATC TTAATCTATC WGRAACTTWA
2941 ATCACMCATT TACCGGAAWA TKTCTGCAAT CTTTATAATT TACARACCCT GATTGTKTCT GGCTGTGAMT
3011 ATTTAGTTAA KTTGCCCAAR ACCTTCTCAA ASCTTAAAAA TTTGCASCAT TTTGACATGA GGGRTACTCC
3081 KAAKTTRAAR AACATGCCCT TARGGATTGG TGARTTGAAA ARTCTACAAA CTCTCTTYMG TAACATTGGC
3151 ATAGCAATAA CCGAGCTTAA GAACTTGCAM AAYCTCCATG GGAAARTTTG TATTGGCGGG CTGGGAAAAA
3221 TGGAAAATGC NGTNGGATGC ACGTTAAGCG AACTTGTCTC A:AAAAAGGT TWAATGARTT ANAAACTGGR
3291 WTKGGGGGTG ATRAATTTAA TGTTTTCCGA AATGGGAACA CTTGAAAAAA NAAGGTCCTC AATGAATTGA
3361 ATGCCTCATA ATGGTAYTCY AAMWAARRRY YYWTARWWAT TWMGKAWRRK GKGTTYATRR TKTTMYRAAW
3431 WAGRGTKTNR KARGTAGGTT TCATCCAATC ACCCAAGTGG GAAAATAGAT GATATTTTCA GGGCYTACTG
3501 ATGAGATGTG GAGAGGTATG ATAGGGTNTC TTGGGGCGGT AGAAGAAATA AGCATCCATT CTTGTAATGA
3571 AATAAGATAT YTGTGGGAAT CAGAAGCAGA GGCAAGTAAG GTTCTTATGA ATTTAAAGAA GTTGGATTTA
3641 GGTGAATGTG AAAATTTGGT GAGTTTAGGG GAGAAAAAGG AGGATAATCA TAATATTAAT AGTGGGAGCA
3711 GCCTAACATC TTTTAGGAGG TTGAATGTAT GGAGATGTAA CAGCTTGGAG CATTGCAGGT GTCCAGATAG
3781 CATGGAGAAT TTGTATATGC ACATGTGTGA TTCAATNACA TCCGTCTCCT TCCCAACAGG AGGAGGACAG
3851 AAGATCAAGT CACTTACCAT CACTGATTGC AAGAAGCTTT CGGAAGAGGA GTTGGGAGGA CGAGAGAGGA
3921 CAAGAGTGCT TATAAACTCA AAAATGCAGA TGCTTGAATC AGTAGATATA CGTAATTGGC CAAATCTGAA
3991 ATCTATCAGT GAATTGAGTT GCTTCATTCA CCTGAACAGA TTATATATAT CAAACTGTCC GAGTRTGGAG
4061 TCATTTCCTG ACCATGAGTT GCCAAATCTC ACCTCCTTAA CAGATCGAAG GAGAGGACAG CGATTTTCGT

RLG1-E169
[Strand]

4131 ACGAACGGTT ACGATTCGAC TGGCCGTCGT TTT

SEQ ID NO: 70

Further Characterization of RG2 Family Members:

Further sequencing of cloned RG2 polynucleotide sequences, as discussed above, identified additional RG2 species, listed below. Additionally, further sequencing of the 5' sections of RG2 sequences listed above resulted in modified and/or new sequence information, also listed below. The AC15 sequences found in the 3' sections of RG2 family have not changed.

Listed below are: four full length species, RG2A, RG2B, RG2C and RG2S; two near complete, but with a gap in the largest intron, RG2D and RG2J; three nearly complete RG2 gene sequences, RG2K, RG2N, and RG2O. The deduced translation products (polypeptides) encoded by these RG2 species are listed below. The polynucleotide sequences do not contain any gaps (as with some of the polynucleotide sequences), because all of the gaps in the sequences are in introns, *i.e.*, there are no gaps in exon, or coding, sequences.

They include: an RG2A polynucleotide sequence (SEQ ID NO:87) and its deduced polypeptide sequence (SEQ ID NO:88); an RG2B polynucleotide sequence (SEQ ID NO:89) and its deduced polypeptide sequence (SEQ ID NO:90); an RG2C polynucleotide sequence (SEQ ID NO:91) and its deduced polypeptide sequence (SEQ ID NO:92); an RG2D polynucleotide sequence (SEQ ID NO:93) and (SEQ ID NO:94), and its deduced polypeptide sequence (SEQ ID NO:95); an RG2E polynucleotide sequence (SEQ ID NO:96) and its deduced polypeptide sequence (SEQ ID NO:97); an RG2F polynucleotide sequence (SEQ ID NO:98) and its deduced polypeptide sequence (SEQ ID NO:99); an RG2G polynucleotide sequence (SEQ ID NO:100) and its deduced polypeptide sequence (SEQ ID NO:101); an RG2H polynucleotide sequence (SEQ ID NO:102) and its deduced polypeptide sequence (SEQ ID NO:103); an RG2I polynucleotide sequence (SEQ ID NO:104) and its deduced polypeptide sequence (SEQ ID NO:105); an RG2J polynucleotide sequence (SEQ ID NO:106) and (SEQ ID NO:107), and its deduced polypeptide sequence (SEQ ID NO:108); an RG2K polynucleotide sequence (SEQ ID NO:109) and (SEQ ID NO:110), and its deduced polypeptide sequence (SEQ ID NO:111); an RG2L polynucleotide sequence (SEQ ID NO:112) and its deduced polypeptide sequence (SEQ ID NO:113); an RG2M polynucleotide sequence (SEQ ID NO:114) and its deduced polypeptide sequence (SEQ ID NO:115); an RG2N polynucleotide sequence (SEQ ID NO:116) and its deduced polypeptide sequence (SEQ ID NO:117); an RG2O

polynucleotide sequence (SEQ ID NO:118) and its deduced polypeptide sequence (SEQ ID NO:119); an RG2P polynucleotide sequence (SEQ ID NO:120) and its deduced polypeptide sequence (SEQ ID NO:121); an RG2Q polynucleotide sequence (SEQ ID NO:122) and its deduced polypeptide sequence (SEQ ID NO:123); RG2S polynucleotide sequence (SEQ ID NO:124) and its deduced polypeptide sequence (SEQ ID NO:125); an RG2T polynucleotide sequence (SEQ ID NO:126) and its deduced polypeptide sequence (SEQ ID NO:127); an RG2U polynucleotide sequence (SEQ ID NO:128) and its deduced polypeptide sequence (SEQ ID NO:129); and RG2V polynucleotide sequence (SEQ ID NO:130) and its deduced polypeptide sequence (SEQ ID NO:131); and, an RG2W polynucleotide sequence (SEQ ID NO:132) and its deduced polypeptide sequence (SEQ ID NO:133).

Characterization of New RG Family Groups and RG Species:

Further BAC insert characterization and sequencing, as discussed above, identified new RG polynucleotide sequences. The new sequences were characterized as belonging to new RG families; designated RG5 and RG7. These RG polynucleotides sequences, and their predicted translation products (the polypeptides which are encoded by these sequences) are summarized and listed below.

Identified and listed below is an RG5 family member, designated as the RG5 polynucleotide sequence set forth in SEQ ID NO:134, and its deduced polypeptide sequence (SEQ ID NO:135). This sequence contains an NBS region sequence.

Also identified and listed below is an RG7 family member, designated as the RG7 polynucleotide sequence set forth in SEQ ID NO:136. No deduced polypeptide sequence is given for the new RG7 family member as this sequence appears to be a pseudogene.

RG2A polynucleotide sequence (SEQ ID NO:87)

AAAGTTCATATCCAAGCTTGCCCTCCAACTCTAGCTCCTTCAATGGCACC
TCCTTCTCTTCAAAAGCACACAAGAACACTTTCAAGCTCAACCACACTCA
CACAAGCTCTAGAACGAGGGTTAGGGCACATTTAGGGTTTTGCTCTCTGG
AAATGGTGTCTAAAAGTGAGGCCATAATGTTCCTTATATAAGGCTCACTC
CCACAATTAGGCTTTCAATCTGAACGTANTACGCCCAGTGTACACTATGG
TACGCCCAACGTACTCGGTAGTCTCCGCGTCAANAATACACTCATGAGTA

CGCGCAACGTACTTTCCCTTACGCCCAGCGTACTCAAAAGCCAAACATTC
TTTTCAAGGACTAATTTTGACAACTTGAGGAAAGAAAAGGATCAAAGANA
TATACTTGAATTCCGGGATGTTACAATGAAGTTGANACCTTGGCTAAAAA
ATTAAATTGGTTGTGGAAGCCGTTGGCTGAGCAAGCAACAAGGGTAAAAT
TCGTAATCTACAAATGGTGTTATTTTCTATTTCTTCTTATTATTTTACTT
GATTTACGGGTAGTTTTTTTTTCTTACAAAAAATATTAAAGTTGATAAAG
TATAGCCACTAAAATTGACTTTTTCCAAAACATAATGTCAAATGGTGCGT
ATATGTATCATGTTGTATTANATAATGAATATGATGATNCTGTTCTATTT
AANCCGAAAAAATTATCTAATGATTTTATATTGGAAAACAAAGTTGTGAT
TTTTNGCATAATATAATCAAATCCNCTTTTGTNTGGGAGGTGGATAAATG
TGGTAAATTTANAACAAGTGTTTTNACNTTGAAGGGTNTGGAAAGGTTGA
AAAAAGTTAAAATGATAAAATGTTTACACAAATGTTGTATCCGACTGAAT
ATNATGTTTAAGGATNATTGTATTAAATTGTTGATATATAGTAAGCATAA
ATATTTAGAATTGTGACTTAAATTTATAAGTTATNCNAACTGGATTGAAA
CATTTTTGATATANATTAGGAATGAAAATGAGCAACCCTAACATACTTAT
CTTTGGTAGTTTGGTTATTATATTTTTATTANAATATAGAANCATCCCTT
TATTTTAAACCCATATTGTGGACGGACTTGAATAAATGGGAAAAATGTAC
CTTGCTATTTAGCACAAAAAAATTATAAAAATGTACATTGCTATTTAGCA
CAAACAAAAAAAAAAAACTTATCCTTTTTGCATTAGGTCACAAAGAAATA
TAAAATGGGAAATGTGTTGCTATTTAATGCACTAAAAGAAACTATTTTGC
CTTTATTAAACCGGGTAAACCAATAGAAAAATGGAAGTACATTGTCATTT
AGCATGAAAAAAAATAACTTTCCATTTTTTGCATCCGGTCACAATAATAG
AAAAATGAAAGTACGTTGCTATTTAGCGAAACTAACTTCCTTTTTTCTTT
TTGGCATCGTATCATAAAATATAGACTAAAATACGTTAGTTTTACATTTT
TAATACATTGAAATGTCTAATCCACATGTTATTCTATAAAAAGGGAAATG
TAATTTACTTATTCTTTGATTCTTTGGCTTCTTTTTAGTACCCAAAACAT
CCCTCTATCCATCTATTCCAACTAAAATAATGAAAACTATATTCCTTCCA
TTGTAGGGATGTTATAAATTTTGTAATTGTTTTATGCAAAAAAGTGTTT
TTTGTTAACTAGATTAACGAGATTCATTTTTCAGCATTTTAGGAGAAGTT
CATCCATCTTTTGGATATGAAGTGCAAGCCAAGTTCTTTAACATGGAATA
TGAGGTCCCTATATGCTCAAAAAATAGCAAATGAGAAATTTTTTAAATTG
GATCCCCATAAAAGAAAATTTGTTAATGGTTGTTTTAATATTGGTCAATG
TGTCCACCGGATGAGCATAATACTAGTTTATAAGGGGTAAAGGTGGGTTT
GGTGGGCCCATTTATCTTTATTATTTCTAAAAGTCAGAATTAAGTAAAAA
AAATTATAAGATAAATACCATAAGGATAAAAAATCATTTTATTTGGACCA
AAGACCAAAGTTGTTAAGGGGCTGTTTGTTTTTTTTGTGAAGAGCTGTGC
AACCACTTTTGTCTGCGCCGCACAGACAACGTGCAGACATATGCCCTCGC
AGAGTGTTTGTTTTTTGAAAGTGCGCAGACCAAAAAAACGTCTGCGCGAG
GTCATCCTGGCGCATATATGTGTCACTGTCTTCAAAGGTCTTCAGACCTC
ATTTTAACCAAAAAAAAAAAAGACCACCGGTTTTTTTTTTTTTTTNTTC
TTTCTCTTGTAGCTGAAAATGCATTTTTAATCTTTATGACATGAAATTAA
GTTTGAAAAATTAATTTATTTCAACAGCTGTAGACGTTAAAAACAAACAG
TCTTCTTGTTGCAGACTGTGGACATTTGGTCCACCTCTTCTACCGCAGAG

ACTTGCAGATGTGGTCCGCAGACTGCAGACATTTTGGCTTCAAATAAACA
AACATCACCTAATTTGACTACACCACACGGACCTCCAATGTAACAAAAAA
AAGGTTGAAACAAAGTTGCCTATTTCTCCATATCCAGGGGCCATTTATGT
AAGAGTTATCTAAATTTTAGTTCGGTAGATCAGTTCTCACATTTTAACCG
GGTAAAGTGTATGTGTGTACGCGCGCACCTGAAAGGTTTGAANGTAACTT
CCAAACTGAANCAANAATCGATATGAAGTATCAAGTTAGAGGTTCAATTG
GTGAAGGAATCAGCTGGAGGTTGGGGAATCGAGCTTCCACTATTAAGGTA
AAATCCATAACCCTAAATGTTGGTACGCTCATATATCAAATTGCGTGTTT
TGTTGAATGAAAAAAGCATGCTCAAAAAACCAGTGTAAGGCACGGTATAT
GACATATTTATAGTTACTGATAACAAATTATGATAATTTTGGGTTTACGT
AAGTTAGGATTCGTACTTCAACCAAATGTAATAGTTTTTGTGAGTCTATC
TATGTATTTGGGGAATCACATTAGCAACGGGATTGTACTAGTAATTCGAA
AAAGTCTTTTAAATAATTTTTCTGTTTATAATTTATGAATAGTTTTAGCG
ACATCTAATATTAAATAGAATGTATCTGATATTGAATTAATGTCCTTAAT
GTGAACATAGACCTTTTCCATTTACTAATGCCTAATTATTAGTTTCTAAT
CAATAAATTTTAATTTCTGTTTTATGCTTCTAAGACAATAAAAATCCATG
ATTTACCTTTAAATATTAACAAAAATGACCATAAATAAATAAAAAATTAG
GATACCAAACCCCCCCGCCATGCCCAATGTCTAAATATTCTTGATGCTTT
TGCTTTTCCCTCTTTTCCTTGTTAGTCTATTATTCTGGAGAGTTTGAGAG
AGTTTCATACAAGAAAATTTCAAGAAGAAAGCAAAGGTCCAGGTATTCTC
TTTTCTTAATTATGTATTAACTTACAAGCATTTTTTACACGATCCATGGT
TTTTTGTGTATGTTTTTCAAATTGAAACTAGATTGGGACTTTTGCCCTTG
ATGATTCATAAGATATTGCATGGAGTTGAGATTGTGTAAGAAAAGTGGTG
AATAGAAAGAGCAAGTGAATCCAGATATAGTATTGGTAATATATGATGAT
GAGATAGAGATATGTTAAAACTGGCTAGAAAATTGTTTTAATTTGAAATT
TAGGTTGTTGAATTTGAAAGATACCAAGCTAATAACTAATTAGTTATGCT
AAATAGTTATAAAGAACAACAAACTCGTAGTTTTTTTTTCATGATTTTCA
ACCTCTTCGTACCAAACTAAATTATAACAAAATTGAATATCATTCTCTGC
AATCAATTTTAACTTTTGTTATTATCATCATGTCTAAAATTGCCACAAGT
TTATTTTCATAGTCATATTGGATTATGAAAGGACTATTTTTACCAATTAC
ATCTTTACTTTATGGCCAAAGCTAATACAATCCGACTAAACTAAAGGATT
CTAGGATGCATATAGTTTGCTCCCCGATTATAGATTTCTATCTAATTTGT
CTATTGTACTAATTTAGGTGCCACCACAAGTAAATTCCTGAAATGGATGT
CGTTAATGCCATTCTTAAACCAGTTGTCGAGACTCTCATGGTACCCGTTA
AGAAACACATAGGGTACCTCATTTCCTGCAGGCAATATATGAGGGAAATG
GGTATCAAAATGAGGGGATTGAATGCTACAAGACTTGGTGTCGAAGAGCA
CGTGAACCGGAACATAAGCAACCAGCTTGAGGTTCCAGCCCAAGTCAGGG
GTTGGTTTGAAGAAGTAGGAAAGATCAATGCAAAAGTGGAAAATTTCCCT
AGCGATGTTGGCAGTTGTTTCAATCTTAAGGTTAGACACGGGGTCGGAAA
GAGAGCCTCCAAGATAATTGAGGACATCGACAGTGTCATGAGAGAACACT
CTATCATCATTTGGAATGATCATTCCATTCCTTTAGGAAGAATTGATTCC
ACGAAAGCATCCACCTCAATACCATCAACCGATCATCATGATGAGTTCCA
GTCAAGAGAGCAAACTTTCACAGAAGCACTAAACGCACTCGATCCTAACC

ACAAATCCCACATGATAGCCTTATGGGGAATGGGCGGAGTGGGGAAGACG
ACAATGATGCATCGGCTCAAAAAGGTTGTGAAAGAAAAGAAAATGTTTAA
TTTTATAATTGAGGCGGTTGTAGGGGAAAAAACAGACCCCATTGCTATTC
AATCAGCTGTAGCAGATTACCTAGGTATAGAGCTCAATGAAAAAACTAAA
CCAGCAAGAACTGAGAAGCTTCGGAAATGGTTTGTGGACAATTCTGGTGG
TAAGAAGATCCTAGTCATACTCGACGATGTATGGCAGTTTGTGGATCTGA
ATGATATTGGTTTAAGTCCTTTACCAAATCAAGGTGTCGACTTCAAGGTG
TTGTTGACATCACGAGACAAAGATGTTTGCACTGAGATGGGAGCTGAAGT
TAATTCAACTTTTAATGTGAAAATGTTAATAGAAACAGAAGCACAAAGTT
TATTCCACCAATTTATAGAAATTTCGGATGATGTTGATCCTGAGCTCCAT
AATATAGGAGTGAATATTGTAAGGAAGTGTGGGGGTCTACCCATTGCCAT
AAAAACCATGGCGTGTACTCTTAGAGGAAAAAGCAAGGATGCATGGAAGA
ATGCACTTCTTCGTTTAGAGCACTATGACATTGAAAATATTGTTAATGGA
GTTTTTAAAATGAGTTACGACAATCTCCAAGATGAGGAGACTAAATCCAC
CTTTTTGCTTTGTGGAATGTATCCCGAAGACTTTGATATTCTTACCGAGG
AGTTGGTGAGGTATGGATGGGGGTTGAAATTATTTAAAAAAGTGTATACT
ATAGGAGAAGCAAGAACCAGGCTCAACACATGCATTGAGCGGCTCATTCA
TACAAATTTGTTGATGGAAGTTGATGATGTTAGGTGCATCAAGATGCATG
ATCTTGTTCGTGCTTTTGTTTTGGATATGTATTCTAAAGTCGAGCATGCT
TCCATTGTCAACCATAGTAATACACTAGAGTGGCATGCAGATAATATGCA
CGACTCTTGTAAAAGACTTTCATTAACATGCAAGGGTATGTCTAAGTTTC
CTACAGACCTGAAGTTTCCAAACCTCTCCATTTTGAAACTTATGCATGAA
GATATATCATTGAGGTTTCCCAAAAACTTTTATGAAGAAATGGAGAAGCT
TGAGGTTATATCCTATGATAAAATGAAATATCCATTGCTTCCCTCATCAC
CTCAATGTTCCGTCAACCTTCGCGTGTTTCATCTACATAAATGCTCGTTA
GTGATGTTTGACTGCTCTTGTATTGGAAATCTGTCGAATCTAGAAGTGCT
TAGCTTTGCTGATTCTGCCATTGACCGGTTGCCTTCCACAATCGGAAAGT
TGAAGAAGCTAAGGCTACTGGATTTGACGAATTGTTATGGTGTTCGTATA
GATAATGGTGTCTTAAAAAAATTGGTCAAACTGGAGGAGCTCTATATGAC
AGTGGTTGATCGAGGTCGAAAGGCGATTAGCCTCACAGATGATAACTGCA
AGGAGATGGCAGAGCGTTCAAAAGATATTTATGCATTAGAACTTGAGTTC
TTTGAAAACGATGCTCAACCAAAGAATATGTCATTTGAGAAGCTACAACG
ATTCCAGATCTCAGTGGGGCGCTATTTATATGGAGATTCCATAAAGAGTA
GGCACTCGTATGAAAACACATTGAAGTTGGTTCTTGAAAAAGGTGAATTA
TTGGAAGCTCGAATGAACGAGTTGTTTAAGAAAACAGAGGTGTTATGTTT
AAGTGTGGGAGATATGAATGATCTTGAAGATATTGAGGTTAAGTCATCCT
CACAACTTCTTCAATCTTCTTCGTTCAACAATTTAAGAGTCCTTGTCGTT
TCAAAGTGTGCAGAGTTGAAACACTTCTTCACACCTGGTGTTGCAAACAC
TTTAAAAAAGCTTGAGCATCTTGAAGTTTACAAATGTGATAATATGGAAG
AACTCATACGTAGCAGGGGTAGTGAAGAAGAGACGATTACATTCCCCAAG
CTGAAGTTTTTATCTTTGTGTGGGCTACCAAAGCTATCGGGTTTGTGCGA
TAATGTCAAAATAATTGAGCTACCACAACTCATGGAGTTGGAACTTGACG
ACATTCCAGGTTTCACAAGCATATATCCCATGAAAAAGTTTGAAACATTT

AGTTTGTTGAAGGAAGAGGTAAATATAAATTTTTAATGCTAATACATTAC
AAAGGATCTTTTCAGTTAAATCTTTCAAAATATATTGTAATTTGATTGTA
TGGGGTATTATTGTTGGATGGGACTATTAATAAATGATTATCTTGCAGGT
TCTGATTCCTAAGTTAGAGAAACTGCATGTTAGTAGTATGTGGAATCTGA
AGGAGATATGGCCTTGCGAATTTAATATGAGTGAGGAAGTTAAGTTCAGA
GAGATTAAAGTGAGTAACTGTGATAAGCTTGTGAATTTGTTTCCGCACAA
GCCCATATCTCTGCTGCATCATCTTGAAGAGCTTAAAGTCAAGAATTGTG
GTTCCATTGAATCGTTATTCAACATCCATTTGGATTGTGTTGGTGCAACT
GGAGATGAATACAACAACAGTGGTGTAAGAATTATTAAAGTGATCAGTTG
TGATAAGCTTGTGAATCTCTTTCCACACAATCCCATGTCTATACTGCATC
ATCTTGAAGAGCTTGAAGTCGAGAATTGTGGTTCCATTGAATCGTTATTC
AACATTGACTTGGATTGTGCTGGTGCAATTGGGCAAGAAGACAACAGCAT
CAGCTTAAGAAACATCAAAGTGGAGAATTTAGGGAAGCTAAGAGAGGTGT
GGAGGATAAAAGGTGGAGATAACTCTCGTCCCCTTGTTCATGGCTTTCAA
TCTGTTGAAAGCATAAGGGTTACAAAATGTAAGAAGTTTAGAAATGTATT
CACACCTACCACCACAAATTTTAATCTGGGGGCACTTTTGGAGATTTCAA
TAGATGACTGCGGAGAAAACAGGGGAAATGACGAATCGGAAGAGAGTAGC
CATGAGCAAGAGCAGGTAAGGATTTCAATTTCACTGTCTTAATTAATGAT
TAAGCTCCTGCTTTTTGAATAAAAAAGGGACAAACCATTTCATGACTTAA
TGTAGCAATACAAGTCATGTATAAGAGTGACCAACTCTTTTTTATTTATA
AAATGACTACAAAATATTTTTTTTCATTAGAGATCATGTATAAATGTGAC
TAATTTTTCATCACCTAACTTTAGTTGATAAATCTTTATAAATGTCACTA
GTTACTTTTCAGTAAAATAACAAATTTAATAAATTATCAACAAAAAGCAT
CAACTAAAAAAATCCCACAACCCGTAATAATTTAAAATAAAAGGATTTAA
CATCTAATACGAACAATTTTTTTTCTAAACATGATTTGGACCAAATATCA
CCAGCAACTCAAGTTTGGAATCGATTCAGCTTAAAACTTGACCAGCATAA
TTAGATAGATGAGAGTTGAAGCTAAAGTGCCTATATAAGTTCGTTTCATC
TTTTTTCTTGATCTTGATAGCAAGTTGAATGATTTTCTTCTTCAAAATTG
ATAAAAATCTACATTATAAAGAGACTAGCTTGAAAAAAAATGGTCTAGGT
GGGTCTTGGGTTCTGGTAGATGAAGATGGAAGGGGAGAGTAGATTTCAAA
GACACAACACATCCTTCATTTTATTTATTTATTATTATTATTATTTTTTG
ATATCTTGCTCATATTTGTTACAGATATGTGAGGTCTATTAATCTTTTTA
AATATATAAAAAAATAAATAACATAAATGAGAAAATTAAATAAAGAATAA
ATTAATAAGGGCACAATAGTCTTTTTAGGTAAGACAAGGACCAAACACGC
AACAAAAATAAACAGTAGGGACCATCCGATTTAAAAAAAATAATTAGGGA
CCAAAAACATAAATTCCCCCAAACCATAGGGACCATTCATGTAATTTACT
CTTACTTTTCGTTTTGTTCATATTTGGGTAACTATTTTTTTTGTACACAT
CTAGGTAACGAACTTGTTGAAGTGTTCCCATTTAGGATGTGACCTACTAC
AACCGATCATAATAGTCATATGTGAACACTTCCAACAACTTTATTACTTA
GGTGTGTACAAAAAAACAATAGTTACCATGATGTGAACATACTGAAAAAT
TAATTACCTTAGCAAGTTATTTTCCCATTTAGGTTGTATGGAAACAGTTC
CGTGAGACCGTGACTTGGATGGTAGATAAATTTAGTAAACTTAACCCTTC
AATTAACCTACCTTTTTCTTATTAACTCAATTTCAACCTAAATTCTGATT

CTTGTTTGAAAGTAAGTTGCATCTTTATTTTTGTATTATCTTGTTGCATA
GGATCCTTAGCATCTTTTAATAATTTATTTGAAGGTGAAAGATCCAACTA
TTTTTAATCTGTTGGCATTTTCCATCATTTGCAACTGTTTCTTGAAAAAA
AAATACCTAAAATCAAAATAACCATTTTCAAATCCAAAATTATAAGAGAG
AATTGTAAATGGACATGGAATCATAAATCATTAACACAGTTCAGTAAACA
AGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTATCAGAGAAA
GAGACATTACAAGAAGCCACTGACAGTATTTCTAATGTTGTATTCCCATC
CTGTCTCATGCACTCTTTTCATAACCTCCAGAAACTTATATTGAACAGAG
TTAAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTGAGAGTCCAACAAGT
AGAGAATTGGTAACAACTCACCATAACCAACAACAACCTATTATACTTCC
CAACCTCCAGGAATTGATTCTATGGAATATGGACAACATGAGTCATGTGT
GGAAGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCA
GAATCCCCATTCCACAACCTCACAACCATAAAAATTATGTATTGCAAAAG
CATTAAGTACTTGTTTTCGCCTCTCATGGCAGAACTTCTTTCCAACCTAA
AGCATATCAAGATAAGAGAGTGTGATGGTATTGGAGAAGTTGTTTCAAAC
AGAGATGATGAGGATGAAGAAATGACTACATTTACATCTACCCACACAAC
CACCACTTTGTTCCCTAGTCTTGATTCTCTCACTCTAAGTTTCCTGGAGA
ATCTGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGGGGAGCAATGAA
ATATCTTTCAATAATACCACTGCAACTACTGCTGTTCTTGATCAATTTGA
GGTATGCTTTGTACATATTCAATTATTTATTTAATTTCCTTTTTTATTTG
CAATATTCTATAAATAATACATTTTATACCCACTATACTAAGATAATAAT
TACCTAGAGGGATGGATGCTATGACACAGCTGCTACACTTCAGAAACTCT
AGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTTTTGAT
GGGTAATATAGGCAATTTAAGTTTTATTTCTGTTAAAGCAGTATTTAGCA
AGTACTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTGAAAATCTGGTC
ATTGTACCCAGAATTTAGTTAAATGTAACATTTTAGATATCAGGGGTCAT
CAGGTGACAGATATTGTAGAATAGAACAATATATAATATCACCCAAAACT
ATTTTTTCTAAGGTTATTCTGTTAAATATGTGCTTTCTTGTTTTCATNGA
ATTNGCATTCGTATATTTTAGGTGTTAAAGTGATTTTNTCTTCAATAAAT
CCCGAAATTAATTAAAAAAAAAAAAAACAAAAGTACATTTTTGATGTGGAG
AGCACTGGTATCACTTAGTATATAAAAAGCTTGATTTTGAATTAACTTTC
TTATACAAAAGTTGTGTATATAGTTTAATTAGTTTACATCATTTTTCCA
TGTGGTGTTGCAGTTGTCTGAAGCAGGTGGTGTTTCTTGGAGCTTATGCC
AATACGCTAGAGAGATGAGAATAGAATTCTGCAATGCATTGTCAAGTGTA
ATTCCATGTTATGCAGCAGGACAAATGCAAAAGCTTCAAGTGCTGACAGT
AAGTGATTGCAAAGGGATGAAGGAGGTATTTGAAACTCAATTAAGGAGGA
GCAGCAACAAAAACAACAAGAGTGGTGCAGGTGAGGAAGGAATTCCAAGA
GTAAATAACAATGTTATTATGCTTTCTGGTCTGAAGATATTGGAAATCAG
CTTTTGTGGGGGTTTGGAACATATATTCACATTCTCTGCACTTGAAAGCC
TGAGACAGCTCCAAGAGTTAAAGATAACATTTTGCTACGGAATGAAAGTG
ATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGTAAACAACAACAAC
AACAACAATAACGAAGGGGGCATCATCATCATCATCTTCTTCATCTTCTA
AGGAGGTTGTGGTCTTTCCTCGTCTCAAATCCATTGAACTAAATGATGTA

CCAGAGCTGGTAGGATTCTTCTTGGGGAAGAATGAGTTCCGGTTGCCTTC
ATTGGAAGAAGTTACCATCAAGTATTGCTCAAAAATGATGGTGTTTGCAG
CTGGTGGGTCCACAGCTCCCCAACTCAAGTATATACACACAGAATTAGGC
AGACATGCTCTTGATCAAGAATCTGGCCTTAACTTTCATCAGGTATATAT
ATTTCTTTAATTGGCATCATCTAATTAAGAAAGATATCATTCCTGCCAAG
TAAATTTACTTCAAACACATTCACACTGGTTTCAGTCTAAGTTTATGTTG
TTCTAGGAAGGCCAAAATGGGAAAGCAAGATAGGGAAAAATAGTGTATTT
CAGTGGAAAGGGTATTTTAGGTATTTTCTGTCAAAAGTTGTTATTGCAGG
CTTTTTAGTACCTGGAATCGTGTGTGGGAGGAGCATTATTATTCTGATTT
GCTTGTTTCTTTATCATTTTTTCTTAGCCTCTGGAACAGCTAGAAACCCT
TTTAATCTTTTGATTTTCAATGACAAAATTTTTCCTGTTACTACATTTGA
TTGTTGTTCTTCATGGTTCTAAGTGAGTTATTGGCTCATCTGTTACTTCT
TTTGATTGTTATTTCATATCATGTTAGTCACTTGAATCAAGCTTTTCTA
TTTTCAACCAGGGCAAAAGGTCAAAAGTAACCTACTTTATGAGATCAAAA
ACAGCAACCCATCGGATAACTTTTAGTTGGAGTTAATAGTTACAATTACC
ATTGTGATTAATAATTATAATATCTTGTATTAATTCATAAAAATTGGTAC
AGCACATATATGACATTTCAAAGGTTTTTGTTTGACATATATATGCCTCT
GGCGTTTTCTTTATTGGACATGCAGACCTCATTCCAAAGTTTATACGGTG
ACACCTTGGGCCCTGTAACTTCAGAAGGGACAACTTGTTCTTTTCATAAC
TTGATCGAATTATATATGGAATTTAATGATGCTGTTAAAAAGATTATTCC
ATCCAGTGAGTTGCTGCAACTGCAAAAGCTGGAAAAGATTCATGTGACTT
ATTGTAATTGGGTAGAGGAGGTATTTGAAACTGCATTGGAAGCAGCAGGG
AGAAATGGAAATAGTGGAATTGGTTTTGATGAATCGTCACAAACAACTAC
CACTACTCTTGTCAATCTTCCAAACCTCAGAGAAATGAAGTTATGGTATC
TAAATTGTCTGAGGTATATATGGAAGAGCAATCAGTGGACAGCATTTGAG
TTTCCAAACCTAACAAGAGTCGATATATGGGGATGTGATAGGTTAGAACA
TGTATTTACTAGTTCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTAC
GCATATGGAACTGCAGTCAGATAGAGGTCGTGATTGTTCAGGATGCAGAT
GTTTGTGTAGAAGAAGACAAAGAGAAAGAATCTGATGGCAAGACGAATAA
GGAGATACTTGTGTTACCTCGTCTAAAGTCCTTGATATTAAAACACCTTC
CAWGTCTTAAGGGGTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATTA
TTGGATACYTTGGAAATCTACRAATGCCCAGCAATAACCACCTTCACCAA
GGGAAATTCCRCTACTCCACAGCTAAAAGAAATTGAAACAMATTTTGGCT
TCTTTTATGCTGCAGGGGAAAAAGACATCAACTCCTCTATTATAAAGATC
AAACAACAGGTAAACCAGATCTTTGTTGCTTNNATAATTCTTAAACNACA
TNTGAAAAGCTTCATGCAAGTTTTTTTNGTTATATNGTCAAAAACCGCAA
CCTACATTTTCAGCTTTANATTTATGTACTTTATGCAGGATTTCAAACAA
GACTCTGATTAATGTGAAGTGAATATTAAAGGTAAATTATATTTTCATGT
TCCTAGTNGCCTATTAATTAAAGGCCTTTTAGTTCGNGATTTTTGGATGT
ATTCTTCATGATGATGTCAATCTTCTAATACCCCATTCATTGTTTGGTTG
AATGTTGACTCTATGTCAGGATGAATATTCAAGGGAAGAATTGTTCATCA
TATGAAGGACATTAAAGAACATGGATGCTCTGAAGATGTTGGGAACACA

RG2A deduced polypeptide sequence (SEQ ID NO:88)

MDVVNAILKPVVETLMVPVKKHIGYLISCRQYMREMGIKMRGLNATRLGVEEHVN
RNISNQLEVPAQVRGWFEEVGKINAKVENFPSDVGSCFNLKVRHGVGKRASKIIEDI
DSVMREHSIITWNDHSIPLGRIDSTKASTSIPSTDHHDEFQSREQTFTEALNALDPNHK
SHMIALWGMGGVGKTTMMHRLKKVVKEKKMFNFIIEAVVGEKTDPIAIQSAVADY
LGIELNEKTKPARTEKLRKWFVDNSGGKKILVILDDVWQFVDLNDIGLSPLPNQGV
DFKVLLTSRDKDVCTEMGAEVNSTFNVKMLIETEAQSLFHQFIEISDDVDPELHNIG
VNIVRKCGGLPIAIKTMACTLRGKSKDAWKNALLRLEHYDIENIVNGVFKMSYDNL
QDEETKSTFLLCGMYPEDFDILTEELVRYGWGLKLFKKVYTIGEARTRLNTCIERLI
HTNLLMEVDDVRCIKMHDLVRAFVLDMYSKVEHASIVNHSNTLEWHADNMHDSC
KRLSLTCKGMSKFPTDLKFPNLSILKLMHEDISLRFPKNFYEEMEKLEVISYDKMKY
PLLPSSPQCSVNLRVFHLHKCSLVMFDCSCIGNLSNLEVLSFADSAIDRLPSTIGKLK
KLRLLDLTNCYGVRIDNGVLKKLVKLEELYMTVVDRGRKAISLTDDNCKEMAERS
KDIYALELEFFENDAQPKNMSFEKLQRFQISVGRYLYGDSIKSRHSYENTLKLVLEK
GELLEARMNELFKKTEVLCLSVGDMNDLEDIEVKSSSQLLQSSSFNNLRVLVVSKC
AELKHFFTPGVANTLKKLEHLEVYKCDNMEELIRSRGSEEETITFPKLKFLSLCGLP
KLSGLCDNVKIIELPQLMELELDDIPGFTSIYPMKKFETFSLLKEEVLIPKLEKLHVSS
MWNLKEIWPCEFNMSEEVKFREIKVSNCDKLVNLFPHKPISLLHHLEELKVKNCGSI
ESLFNIHLDCVGATGDEYNNSGVRIIKVISCDKLVNLFPHNPMSILHHLEELEVENC
GSIESLFNIDLDCAGAIGQEDNSISLRNIKVENLGKLREVWRIKGGDNSRPLVHGFQS
VESIRVTKCKKFRNVFTPTTTNFNLGALLEISIDDCGENRGNDESEESSHEQEQIEILS
EKETLQEATDSISNVVFPSCLMHSFHNLQKLILNRVKGVEVVFEIESESPTSRELVTT
HHNQQQPIILPNLQELILWNMDNMSHVWKCSNWNKFFTLPKQQSESPFHNLTTIKI
MYCKSIKYLFSPLMAELLSNLKHIKIRECDGIGEVVSNRDDEDEEMTTFTSTHTTTT
LFPSLDSLTLSFLENLKCIGGGGAKDEGSNEISFNNTTATTAVLDQFELSEAGGVSW
SLCQYAREMRIEFCNALSSVIPCYAAGQMQKLQVLTVSDCKGMKEVFETQLRRSSN
KNNKSGAGEEGIPRVNNNVIMLSGLKILEISFCGGLEHIFTFSALESLRQLQELKITFC
YGMKVIVKKEEDEYGEQ.TTTTTTITKGASSSSSSSSKEVVVFPRLKSIELNDVPELV
GFFLGKNEFRLPSLEEVTIKYCSKMMVFAAGGSTAPQLKYIHTELGRHALDQESGL
NFHQTSFQSLYGDTLGPVTSEGTTCSFHNLIELYMEFNDAVKKIIPSSELLQLQKLEK
IHVTYCNWVEEVFETALEAAGRNGNSGIGFDESSQTTTTTLVNLPNLREMKLWYL
NCLRYIWKSNQWTAFEFPNLTRVDIWGCDRLEHVFTSSMVGSLLQLQELRIWNCSQ
IEVVIVQDADVCVEEDKEKESDGKTNKEILVLPRLKSLILKHLPCLKGFSLGKEDFSF
PLLDTLEIYKCPAITTFTKGNSTTPQLKEIETHFGFFYAAGEKDINSSIIKIKQQDFKQ
DSD.CEVNIK

RG2B polynucleotide sequence (SEQ ID NO:89)

TTTTTTAAGATCAGGGATTCAAATTCAGCCCTAGTGATTACAATTGTGTC
TAAACTTTCCCATACCTTCACATTATTGTAAGTATACTTTCTCAGTTTCT
CTCTTGGAAGCTTCCTTGGTATTTTAACTCGTGTTCTAATATTTAACTCT
GATAGTTATTTTGGCCAATCTACTATCTGCATGTCCGGTTATTGAATCCG
AAGGCACTGGAATCTTGGATTCCATTCCGTTGTGTGTTTGGTTGCCAAAT

GAACGGAATTGAATTATGTAAGATTCCTTCAAAATCCATGTTTAGGTATA
TCGTTGTTTCTTGGGATGGATGGTAAAGAACGGAATTTCTCCTGTTCATT
TTTTAATGAAAGACCAAATTGACCTTATAAACCTGTTAAAAAAATTACAT
TCCAGTTTTCTTAACAAACTGAAAATGGTAAAGGAGTGTGATTGAATTCC
AATCTGTTTCCTGTCCAAAACACGTGACGGAATATTACAATTCCTTCAAA
TTTCATTTTCTTAAATTGTTATTCCCTTTCTTACAAAAACAAGGTAAACG
AAACACCCGCTTACTTAATCATACTCCTACATGATGTAAATGAAAAGGGT
ATAAATGGTATTTTATTCACAGGGATGAGTCACCATGGTCATGAAAGAAT
CATTAACCGCCCTTACCCAATTCATGTTTGCCCCTAAAATATGATTTAAA
GTAATATTGGCTTATGGGATTCAAGTTGACTTTTTTGTGGCGAAGAAATA
ATGAAAATCTTCATTTCTAAAGTGTCTTCTACCACTGACATTTTCTAAGA
AAGAACTTGCTAGAAGAAGGTGGGTTGTTTAGTCTTTTTACTCTTTAAAT
GTGAAGACTGTTGAGTTATTATTATTATTTTGCCAACTATGGACAACTTG
TTTAGTTTTTTTTTTCCCCAATATCCATTTATATGCGATTTATTTCTGA
AATAATTTTATCAAAACGCAGGAAACAATGTAGAATAATACTGGTATAAT
TAATTATATAAAGTTATTAGGCTGAAATCTTGAGGCTACTATAATTTAAT
TATCATAATTTGAAAATCATCAAATTGTATTCCATGTATATTTATGTTAT
CAGATAATTAATAATATGTGAGCCACACAAATCCACATCATCAGACACCC
CACCTTATTGTCGGCTACCTCACCACTTGCATGATCCCGACATCTTCCCA
ACCCCACCGACGACTTGGGGTCTCCTTAATATATCAATTATTTTCTGTAA
GTATTTATTTGTGTAAATGTGTAATGTCATTTTACCTTTTTTCTAATATA
TACAGAAACATAAATTTTAAATGAAATTCAACTGCGTTTCATTCTTGCAT
TAAAAAAAAAGACTGTACTGTTGTCAATATTTTACTTATAACCTGATTAA
TTAATTAAAGCGTAATTGCATAATTTGCATTAGGTTGTAATTTTGTGTTT
TATAGGGAGGGTGAGGGTCACCGGGAATCAAAGCACTTATGTAAAAGCAG
GGAAATACAAAAAATTTACTCGAAACAAATTTTATTCAATTTAAGTGAGA
TAATAATGTTCTGATTAGATTATGAGAACTAGGAGATTTAAGTGATATAT
CCCATTTAAAAGAAATTGCATTATTAATTTTGGATCTCTTGATGATGACA
AAATTAACTCGTGACAGGTTATATATCATATACAAAATGAGTGGCTATGC
TTTCGCTTTCCAAAAAGCAATTATAGTTATACTACACCTACAAATTTTAA
AAGGGGTTAAACATATCAAAATACTTGATAAGTAATTATATAAATATGCA
TTTAACCCTCTAAAGAAAATGCTACTAAGCTTGGACCATCTCAGAATTAC
AATCATACCCTTCCCCTCAAAAAAGATTCGTATATATCATGTCATTTGGC
ATTCATTTCTTTTTCACAATTCATAGTTCTATTCTCAAAAAATTCGAGTT
CTCGTATTTGTAAGGAAGATCAGAAGAGACTGTTCACACAGGTACTCTCT
TTTATTTATTGATTCACATTCATATATGTTATTGTTTTCTTGCTTAATGG
TTTCGTCAGTCTAACTGCGCTTGCTGATTTAAATTTCTTCACTTTCTTCC
ACGGATTTTTTAAATATTAGTTTTGTGAATGAACAATTGGTGAAGGAAAG
AAACATGGGAGTCTTTTCTAAAGTAAACCTAGATACTTAGGTTATAAGGG
TATATGCTAAAATGAACTATGCCCATTCACCTTTGCCTTTTCTTTTACTT
TTTAGTTTTTAGAATCCAAGTTTTCATATGTATCTCGATGTGTGAGAAGA
ATAGGCATTAGAAAGGTAAAGGACGTACATAAAATTGATTAATTAGTGAA
TGTTCTTTGATATCATTATTTTTACTCTCATAAAAAGCATATAGATCAAA

CACAAATTGCTACTTGTTAGTGTAACAACTTCGACTTAATAATGTTAATA ATCAAGATTCTCTTGATTTCAACTATTTCTAACCGAACAAGCTCACTAA AAACTCATATTGCTTTGAGTCTGAGTGGTTTATATTTGGGGTTTTACATT TAATTTTTTGTGCATGAATGTGAAAATAGACTGCTTATTGATTCTTTGTG TTTCATTGAGTTGATTTTCATTATTACTACCTTACAAATTGCTCAGTGAT AGATTTCCATTAATTTGCTAATTCGGTTGCTTCTAAATATGTAGGAGCTA CTAAAAGCAAAAATATCGAGCAATGTCGGACCCAACGGGGATTGCTGGTG CCATTATTAACCCAATTGCTCAGACGGCCTTGGTTCCCGTTACGGACCAT GTAGGCTACATGATTTCCTGCAGAAAATATGTGAGGGTCATGCAGATGAA AATGACAGAGTTGAATACCTCAAGAATCAGTGTAGAGGAACACATTAGCC GGAACACAAGAAATCATCTTCAGATTCCATCTCAAACTAAGGAATGGTTG GACCAAGTAGAAGGGATCAGAGCAAATGTGGAAAACTTTCCGATTGATGT CATCACTTGTTGTAGTCTCAGGATCAGGCACAAGCTTGGACAGAAAGCCT TCAAGATAACTGAGCAGATTGAAAGTCTAACGAGACAACTCTCCCTGATC AGTTGGACTGATGATCCAGTTCCTCTAGGAAGAGTTGGTTCCATGAATGC ATCCACCTCTGCATCATTAAGTGATGATTTCCCATCAAGAGAGAAAACTT TTACACAAGCACTAAAAGCACTCGAACCCAACCAAAAATTCCACATGGTA GCCTTGTGTGGGATGGGTGGAGTGGGGAAGACTAGAATGATGCAAAGGCT GAAGAAGGCTGCTGAAGAAAAGAAATTGTTTAATTATATTGTTGGGGCAG TTATAGGGGAAAAGACGGACCCCTTTGCCATTCAAGAAGCTATAGCAGAT TACCTCGGTATACAACTCAATGAAAAAACTAAGCCAGCAAGAGCTGATAA GCTTCGTGAATGGTTCAAAAAGAATTCAGATGGAGGTAAGACTAAGTTCC TCATAGTACTTGACGATGTTTGGCAATTAGTTGATCTTGAAGATATTGGG TTAAGTCCTTTTCCAAATCAAGGTGTCGACTTCAAGGTCTTGTTGACATC ACGAGACTCACAAGTTTGCACTATGATGGGGGTTGAAGCTAATTCAATTA TTAACGTGGGCCTTCTAACTGAAGCAGAAGCTCAAAGTCTGTTCCAACAA TTTGTAGAAACTTCTGAGCCCGAGCTCCAGAAGATAGGAGAGGATATCGT AAGGAAGTGTTGCGGTCTACCTATTGCCATAAAAACCATGGCATGTACTC TTAGAAATAAAAGAAAGGATGCATGGAAGGATGCACTTTCGCGCATAGAG CACTATGACATTCACAATGTTGCGCCCAAAGTCTTTGAAACGAGCTACCA CAATCTCCAAGAAGAGGAGACTAAATCCACTTTTTTAATGTGTGGTTTGT TTCCCGAAGACTTCGATATTCCTACTGAGGAGTTGATGAGGTATGGATGG GGCTTGAAGCTATTTGATAGAGTTTATACGATTAGAGAAGCAAGAACCAG GCTCAACACCTGCATTGAGCGACTGGTGCAGACAAATTTGTTAATTGAAA GTGATGATGTTGGGTGTGTCAAGATGCATGATCTGGTCCGTGCTTTTGTT TTGGGTATGTTTTCTGAAGTCGAGCATGCTTCTATTGTCAACCATGGTAA TATGCCTGGGTGGCCTGATGAAAATGATATGATCGTGCACTCTTGCAAAA GAATTTCATTAACATGCAAGGGTATGATTGAGATTCCAGTAGACCTCAAG TTTCCTAAACTAACGATTTTGAAACTTATGCATGGAGATAAGTCGCTAAG GTTTCCTCAAGACTTTTATGAAGGAATGGAAAAGCTCCATGTTATATCAT ACGATAAAATGAAGTACCCATTGCTTCCTTTGGCACCTCGATGCTCCACC AACATTCGGGTGCTTCATCTCACTGAATGTTCATTAAAGATGTTTGATTG CTCTTCTATCGGAAATCTATCGAATCTGGAAGTGCTGAGCTTTGCAAATT

CTCACATTGAATGGTTACCTTCCACAGTCAGAAATTTAAAGAAGCTAAGG
TTACTTGATCTGAGATTTTGTGATGGTCTCCGTATAGAACAGGGTGTCTT
GAAAAGTTTTGTCAAACTTGAAGAATTTTATATTGGAGATGCATCTGGGT
TTATAGATGATAACTGCAATGAGATGGCAGAGCGTTCTTACAACCTTTCT
GCATTAGAATTCGCGTTCTTTAATAACAAGGCTGAAGTGAAAAATATGTC
ATTTGAGAATCTTGAACGATTCAAGATCTCAGTGGGATGCTCTTTTGATG
AAAATATCAATATGAGTAGCCACTCATACGAAAACATGTTGCAATTGGTG
ACCAACAAAGGTGATGTATTAGACTCTAAACTTAATGGGTTATTTTTGAA
AACAGAGGTGCTTTTTTTAAGTGTGCATGGCATGAATGATCTTGAAGATG
TTGAGGTGAAGTCGACACATCCTACTCAGTCCTCTTCATTCTGCAATTTA
AAAGTTCTTATTATTTCAAAGTGTGTAGAGTTGAGATACCTTTTCAAACT
CAATCTTGCAAACACTTTGTCAAGACTTGAGCATCTAGAAGTTTGTGAAT
GTGAGAATATGGAAGAACTCATACATACTGGAATTGGGGGTTGTGGAGAA
GAGACAATTACTTTCCCTAAGCTGAAGTTTTTATCTTTGAGTCAACTACC
GAAGTTATCAAGTTTGTGCCATAATGTCAACATAATTGGGCTACCACATC
TCGTAGACTTGATACTTAAGGGCATTCCAGGTTTCACAGTCATTTATCCG
CAGAACAAGTTGCGAACATCTAGTTTGTTGAAGGAAGGGGTAGATATATG
TTCTTTATGTTAATACAATTTAAATAATATTTTCAACCAAATTTTCATAA
TATATCTGTAATTTGATTGTATGATGTGTTATTGTTTATATGTGGCTATT
AAGGGATGATTATTTTGCAGGTTGTGATTCCTAAGTTGGAGACACTTCAA
ATTGATGACATGGAGAACTTAGAAGAAATATGGCCTTGTGAACTTAGTGG
AGGTGAGAAAGTTAAGTTGAGAGCGATTAAAGTGAGTAGCTGTGATAAGC
TTGTGAATCTATTTCCGCGCAATCCCATGTCTCTGTTGCATCATCTTGAA
GAGCTTACAGTCGAGAATTGCGGTTCCATTGAGTCGTTATTCAACATTGA
CTTGGATTGTGTCGGTGCAATTGGAGAAGAAGACAACAAGAGCCTCTTAA
GAAGCATCAACGTGGAGAATTTAGGGAAGCTAAGAGAGGTGTGGAGGATA
AAAGGTGCAGATAACTCTCATCTCATCAACGGTTTTCAAGCTGTTGAAAG
CATAAAGATTGAAAAATGTAAGAGGTTTAGAAATATATTCACACCTATCA
CCGCCAATTTTTATCTGGTGGCACTTTTGGAGATTCAGATAGAAGGTTGC
GGAGGAAATCACGAATCAGAAGAGCAGGTAACGCTTTCAATTTCACTTTC
TTAATTAATTAAGGACTAAGCTCCTGTTTTTGAATAATAAAGAGGTGGG
ATGACTAAACTTGGGCATCACAATTGCAACAAAATGTTACAAACCATGAA
ACGTTCAAACCATTTCTTGAATTAAGGTTTCAATACAAGTCATTTAAAAA
TATGGCTTAAATTTTTTTTATATTTATGTATCAACATGATTTTTCATTAG
AGATCATTATTATAATAGTAAGTTTAAAGCAATTTAAATCAGAACTAATT
CTAACTTTAGCTAATAAATCGTTATAAATGTAAATAATTACTTTTTAGTG
AAATAAGCAACGGATTTAATAAGTTAACAACTTAAATGTCATTTCCTAAC
AAAAAAAACTATTTGGTTCAGAAAAACCGTAATTCAAGATAACTAAAATA
AAAATATTTGACATTCACTAAGAGCATTTTTTTTCTAAATATGATTGCA
AATGAATAAAACTTAAATTTATACAGAAAATTCTTTTATATATGTTATAC
AAAATTTACAAATTGAAATTGGATATGTTAATTAACGGTTTATAATTCTG
GTATCACAAAGGGATATATAATAAAATATTATTTTCTGTAGTCATTTGTA
ATTGTACTAGTTTATAACCCGTGGGAACCATGAGTTCTAAAATTAGTTAA

ACTTTCATAATAAAAATTTATAATTATTATTTATTTTAAATAAATTATTA
ATTAAGAGATATATCAAAAATTTAAAGTTATTATAACTTCAAATTTAACA
TATAATTAGAAAATATATGATCATAACTTTCTGCAACTCTTCTTTTGTAT
TAAAATGACCAGAGAAGCTCTTAGTATATTTCTAATCAAAGTCTCAAAMC
TAATGAAGCATATAATTTGTGAAAATCAATTAGCATTAGGTTTTAAGAGT
CACCAAATTCAAAGAATAATCCAATGCTTTCATTACCACTATGGAGAAAA
TATTTTCTTAGTTTAAATGAAATGAAAACAAACATTCAAACTAATTGTTG
CTTATTAAACCAAAGACCCATTACTTAGCCAAGAGTTTAACAAAAAAAAA
TTACATTCATGTATCATTATTCATGACTAGATATATATGAACATGAAGGG
AGTTTTTATAGAAAATATAATCATAGATATTCAACATAACTTCAGGGAAT
TCCTCAAAATAACCAAGTTATTCAAGAAATTACATCCAAGTCAACCAAAG
AGAAGTTTAGCCTAGCATGGCTAAACTCAAGAAACTAAAATAAGGATTAG
AAGTACCAAACATGTAGTAAGAATCACAGTAAAAGATGATGTTGTTCTTG
ATGTTCTTCTAAGTTCTTCAAGTCTCCAGTTGCTCCTAATAATGCAAAGG
AGAGCCATTAAATTCGTATGTATTGATCCCTTCAAAAGCTGCACCAACCT
CCCTTAAATAACACTCAAAGCAAAAATGACAAAATTGCCCCTGAAGGACC
CTATGTGGGTGCCTTGCGCGGGTGGAGCTGCATACGAAAGGTCTTTGGTC
TTTGTGAGGGTGATGTTGTGCGGGATAGCTTGTCGCATGCTTCCGCGCGG
TTCACGCACATGTGCACAGGTGATGCATGGTGTGTGCGTTCTTGAGTTTT
GAGCCTCCGATGCTTAGTCCACTTGGCCCAATTCGAGTCCAATCAGCTTA
TAACCCATTTTTCTTCAAGTTATCTTCAAGTTAAGCCCAATTTGCCTTCT
CCAAATCATCCATAACTTCACAGAATCGCCCGTTCATCTTAATCCCGGAT
GCACAATTATTCTCCCGTCTTCATTTTAAGCAAGATACCACCTTCTTCAT
GCTTCATCCATCAATAGTACACTTCATGTATCATCTCTACTAGTTATTTA
GTCCACAATCCTTGTTGTCCTCCAAATTTAATTATCTCATTTAGTTCCCG
TTCCGCTAGTTTCCTTAAAATTTGCAATTAAGCTCAGAGAAATATTAAGT
ACCCGAAATGGTCATAAAATAACAAAAAGGAAAATATGCATGAAGATTAA
CTAAATGATGAACGAAATATGCTAAAATAGACTATAAAATGAAGTAAATA
AAATGAAATTATCGCACTCCGACCACCCTTATAGGCTTGTAGTCCACCCA
CCCTTCATTCCTTGTACCAATATGGGATGGAAACATCATTAATTAAGCCA
AAAAGCTAACATATAAGGGTTTAGTGACAAAGGTAAGTACTAAAGATGAA
AATAATCCATTTTTCTTGTATATACACAACACACACATAGGGGCAGACGT
AGGATTTCAAAGTACAGATTGTTGGTGGCACATAAGTGTTGCTGGTGACA
TTTTTTTTTTTTTTTACGTAGTGGCACAACAGTAGGAAAAACGAAAAAT
TCGAAATTTTTTACAATTTGTCTAAAAAAAACAGTGGTTGTTGGTGCCAC
TATGGACACCAAAGTTGAACTGCCCCCACGCGCGCACACACACACACACA
CATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAAAGAAAGAAAGAGAGA
GAGAGTTTGGGATGTGATACTTCTTTTAGGAAAATGGAGTTATATCTTTG
ATATTGTATTTTTTTAATGTAATTTATATATTTAATCATTTTAGTTTATA
AGTTTTATTTATTTTGATATGAAAAAAAAAGTCTTTTATACATTGGATTT
AACATAAAAATCCAACAATATTAATCAAAAAGACCAMACATGTGGACAMW
TATGTATATAAWTAATTCACAATAGTCTTTAGGAATAGNATTATATATAT
AATTAATTCTCAATGGTCTTAGGAATAGTAAGTTCTTATATTTCAAACTT

TNGCCACAATTCTTTGKTTACTTWGACACTTYCCTCTCTCTAATTATATA
TATATATATATATATATATATATATATACACACACACACACACACACTAG
ATGTGTGCCCGCGCAAAGCAGTGACGTNNNGGAGAANACTTTCTTAAGCA
TAAATAATTATTATATTTTTTATTGGGTATTATATAATAAAAAATTACAA
CTTTTAAATAAAATATTTATGTTTATACTTTATATTTATATTGCTTGTAT
ACTATTAATATAATAAATTAATATTTATGTCTAATTTATGAAATGTAAAT
TAATTTAAATACATGAATTTAATATTTTTAAAATTTTCAGTTTGCTTCAA
ATTGAGTTTCTTAATTATTTTTTTAATTCANGTATTCAAACTTTTGGTA
AGTATTAAAGAATTATTTATGCATAATTGATTTATACAAAAAACTTTGTA
ACTTATACATCTTAAAATTCAAGATATAACTAACATGTTTTACAATATAT
ATATATATATATATATATATATATATATATATATATATATATATATATAT
TAAAGCGCAAAGGTCATAGGAATAGAATATTTTCTATTATTCTACGTTTT
GCCACAAAAGTTTGAACACTTTGCCACTTTTTGTCCCTCCTTAACCTTTT
CAATGTTTTGCGACAAAAGTTCCAAAACTTTGCCACTTTGATCATTCCTC
AACTTTTCACCGCAATTAGTTTGTGGAGTTGGCAGTTTTGATCCCCCTAA
CTTCGATATTCTCTACTGCTAGCCAAAAAGGGTTCCAGAGTTTCACACTT
TTGGTCCCTGACAGTAACCAAATGTGAGATGTCAAATTTTTGCCACATTA
GTTTGTGGAGTTGTCCCTTTTGGTCCCCCCACATTCGATATTCTANTATA
CGACCTTATTTTTNTCAAATAACAACACGTATATTTAATTACCAATTATA
GAAATAGATATCAAATAAAGTATTTGTAACACTGTGTAAGAACGGTGCTA
CTATAGGTAAAAATAAACATTTCAAAGTACGATATCCTAATTGGAAAAAG
AGTTTTAAAAAAATAACGACTAGGGGCGAGTTTTTTTACAAGTTTGTAT
CAAATCATATCAAAATTTAAGGTGGAACGGTGACCACATTAACCAGAAAT
GTAATTTATTCTTTGATTTTGATAATTTTTAATATTTTGTTGTGATCTAT
GTATTTAAAAGTAAACAACAAAGAACATAATCCAAAACCCTAAATTGCAA
GTCTCGCCCAATTTCTCTATCACTAGTCCTCACTTACGATGGCGTTACGT
CGCTCTCTCACTGCTTACAACCCTTTGTTGCTACTCATTACAATAACGAA
AAGTTGAATATCCATATATTTATTTGGATGTGGAATTGAACGAATCTCGT
CAAAATTTTGATTTTGTTGATGGATTTGAGTAGAAGTTTGGGCAGAACGG
GAATGATGGTCTGCAAGTGGTTATAAACTTGATTCTGAGTTATTACTATA
TATGTAGCCTCTTTACAACGACCAAGGTTTCTTCCAGGTACCATTTGATC
TTTTTAGAACTTAGTTTTCTGAAACACCCTGATTTGGATCAAATATCACC
AACAACTCTTAAAAACTTGATTAATCAATTGTTTTCTTCATCTTGATAAC
AAGTGGAATGATTTTCTACTTAGATTAACTTGAAAAAAAAGGTCCATGTG
CGTCTGGTGGATCTGGTAAATGAAGATGGAAGGGAGAGCTGACTTTAAAG
ACACAAACACGTCACCATATCTCTTATTTTATTTTAAATTTGCTTTTGGT
GTATTTTCTTTTTTCCTATTTCTTTCTTTCTTGATCTCCAGATGGTATGT
GGTGTGGATAATTTACACCTAGAGATTGGGAACGATGGGAAGGGGTCTGT
GATTTATGGCTGGCCGAGTTTTACTTATTAACTCAATTTCAACCTAAATT
CTGATTCTTGTTTGAAAATAAGTTGCATCTTTATTTTTGTATTATCTTGT
TGCATAGGATCCTTAGCATCTTTTAATAATTTATTTGAAGGTGAAAGATC
CAACTATTTTTTAGCTGTTGGCATTTTCCATCATTTGCAACTGTTTCTTG
AAAAAAAAATACCTAAAATAAAAATAACCATTTTCAAATCCAAAATTATA

AGAGAGAATTGTAAATGGACATGGAATCATAAATCATTAACACAGTTCAG
TAAACAAGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTATCA
GAGAAAGAGACATTACAAGAAGCCACTGGCAGTATTTCAAATCTTGTATT
CCCATCCTGTCTCATGCACTCTTTTCATAACCTCCGTGTGCTTACATTGG
ATAATTATGAAGGAGTGGAGGTGGTATTTGAGATAGAGAGTGAGAGTCCA
ACATGTAGAGAATTGGTAACAACTCGCAATAACCAACAACAGCCTATTAT
ACTTCCCTACCTCCAGGATTTGTATCTAAGGAATATGGACAACACGAGTC
ATGTGTGGAAGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAA
CAATCAGAATCCCCATTCCACAACCTCACAACCATAAATATTCTTAAATG
CAAAAGCATTAAGTACTTGTTTTCGCCTCTCATGGCAGAACTTCTTTCCA
ACCTAAAGGATATCCGGATAAGTGAGTGTGATGGTATTAAAGAAGTTGTT
TCAAACAGAGATGATGAGGATGAAGAAATGACTACATTTACATCTACCCA
CACAACCACCACTTTGTTCCCTAGTCTTGATTCTCTCACTCTAAGTTTCC
TGGAGAATCTGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGGGGAGC
AATGAAATATCTTTCAATAATACCACTGCAACTACTGCTGTTCTTGATCA
ATTTGAGGTATGCTTTGTACATATTCAATTATTTATTTAATTTCCTTTTT
TATTTGCAATATTCTATAAATAATACATTTTATACCCACTATACTAAGAT
AATAATTACCTAGAGGGATGGATGCTATGACACAGCTGCTACACTTCAGA
AACTCTAGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCT
TTTGATGGGTAATATAGGCAATTTAAGTTTTATTTCTGTTAAAGCAGTAT
TTAGCAAGTACTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTGAAAAT
CTGGTCATTGTACCCAGAATTTAGTTAAATGTAACATTTTAGATATTAGG
GGACATCAGGTGACAGATATTGTAGAATAGAACAATATATAATATTACCC
AAAACTATTTTTTCTAAGGTTATTCTGTTAAATATGTGCTTTCTTGATTT
CATTGAATTTGCATTCCTATATTTTAGGTGGTAAAGTGATTGTCTCTTCA
ATAAATCCCGAAATTAATTAAAAAAGAAAAAAACAAAAGTAAATTTTTGA
TATGGAGAGCACTGGTATCATTTAGTATATAAAAAAACTAGATTTTGAAT
TAAGTTTCTTATATAAAAGCTGTGTATATAGTTTAATTAGTTTTACATCA
TTTTTCCATGTGGTGTTGCAGTTGTCTGAAGCAGGTGGTGTTTCTTGGAG
TTTATGCCAATACGCTAGAGAGATARAAATAGKTGGATGCTATGCATTGT
CAAGTGTGATTCCATGTTATGCAGCAGGACAAATGCAAAAGCTTCAAGTG
CTGAGAATAGAGTCTTGTGATGGCATGAAGGAGGTATTTGAAACTCAATT
AGGGACGAGCAGCAACAAAAACAACGAGAAGAGTGGTTGCGAGGAAGGAA
TTCCAAGAGTAAATAACAATGTTATTATGCTTCCCAATCTAAAGATATTA
AGTATTGGAAATTGTGGGGGTTTGGAACATATATTCACATTCTCTGCACT
TGAAAGCCTGAGACAGCTCCAAGAGTTAAAGATAAAATTTTGCTACGGAA
TGAAAGTGATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGCAAACA
ACAACAACAACAACGAAGGGGGCATCTTCTTCTTCTTCTTCTTCTTCTTC
TTCTTCTTCTAAGAAGGTTGTGGTCTTTCCTTGTCTAAAGTCCATTGTAT
TGGTCAATCTACCAGAGCTGGTAGGATTCTTCTTGGGGATGAATGAGTTC
CGGTTGCCTTCATTAGATAAACTTAAGATCAAGAAATGCCCAAAAATGAT
GGTGTTTACAGCTGGTGGGTCCACAGCTCCCCAACTCAAGTATATACACA
CAAGATTAGGCAAACATACTCTTGATCAAGAATCTGGCCTTAACTTTCAT

CAGGTATATATATATATTTCTTTAATTGGCATCATCTAATTAAGAAAGATAT CATTCCTGCCAAGTAAATTTACTTCAAACACATTCACACTGGTTTCAGTC TAAGTTTATGTTGTTCTAAGAAGGCCAAAATGGGAAAGCAAGATAGGGAA AAATAGTGTATTTCAGTGGAAAGGGTATTTTAGGCATTTTCTGTCAAAAG TTGTTATTGCAGGCTTTTTAGTACCTGGAATCGTGTGTGGGAGGAGCATT ATTATTCTGATTTGCTTGTTTCTTTATCATTTTTTCTTAGCCTCTCGAAC AGCTAGAAACCCTTTTAATCTTTTGATTTTCAATGACGAAATTTTTCCCT GTTACTCCATTTGATTGTTGTTCTTCATGGTTCTAAGTGAGTTATTGGCT CATCTGTTACTTCTTTTGATTGTTATTTTCATATCATGTTGTCCTTTGAA TCAAGCTTTTCCATTTTCAACCAGGGCAAAAGGTCAAAAGTAACCTACTT TATGAGATCAAAAACAGCAACCCATCGGATAACTTTTAGTTGGAGTTAAT AGTTACAATTACCATTGTGATTAATAATTATAATATCTTGTATTAATTCA TAAAAATTGGTACAGCACATATATGACATTTCAAAGGTTTTTGTTTGACA TATATATGCCTCTGGCGTTTTCTTTATTGGACTTGCAGACCTCATTCCAA AGTTTATACGGTGACACCTTGGGCCCTGCTACTTCAGAAGGGACAACTTG GTCTTTTCATAACTTTATCGAATTAGATGTGGAAGGTAATCATGATGTTA AAAAGATTATTCCATCCAGTGAGTTGCTGCAACTGCAAAAGCTGGAAAAG ATTAATGTAAGGTGGTGTAAAAGGGTAGAGGAGGTATTTGAAACTGCATT GGAAGCAGCAGGGAGAAATGGAAATAGTGGAATTGGTTTTGATGAATCGT CACAAACAACTACCACTACTCTTGTCAATCTTCCAAACCTTAGAGAAATG AACTTATGGGGTCTAGATTGTCTGAGGTATATATGGAAGAGCAATCAGTG GACAGCATTTGAGTTTCCAAACCTAACAAGAGTTGATATCTATAAATGTA AAAGGTTAGAACATGTATTTACTAGTTCCATGGTTGGTAGTCTATCGCAA CTCCAAGAGCTACATATATCCAACTGCAGTGAGATGGAGGAGGTGATTGT TAAGGATGCAGATGATTCTGTAGAAGAAGACAAAGAGAAAGAATCTGATG GGGAGACGAATAAGGAGATACTTGTGTTACCTCGTCTAAACTCCTTGATA TTAAGAGAACTTCCATGTCTTAAGGGGTTTAGCTTGGGGAAGGAGGATTT TTCATTCCCATTATTGGATACTTTAAGAATTGAGGAATGCCCAGCAATAA CCACCTTCACCAAGGGAAATTCCGCTACTCCACAGCTAAAAGAAATTGAA ACACATTTTGGCTCGTTTTGTGCTGCAGGGGAAAAAGACATCAACTCTCT TATAAAGATCAAACAACAGGTAAATCAGATCTTTGTTGCTTTAATAATTC TTAAACTACATTTGAAAAGCTTCATGCAAGTTTTTTTGTTATATTGTCA AAAACCGCAACCTACATTTTCAGCTTTATATTTATGTACTTTATGCAGGA GTTCAAACAAGACTCTGATTAATGTGAAGTAAATACTAAAGGTAAATTAT ATTTTCATGTTCCTAGTTGCCTATTAATTAATTGCCTTTTAGTTCATGAT TTTTGGATGCATTCTTCATGATGATGTCAATCTTCTAATACCCCATTCAT TGTTTGGTTGAATGTTGACTCTATGTCTTGATGAATATTCAAGGGAAGAA TTGTTCATCATATGAAGGACATTAAAGAAGAACATGGATGCTATGAAGAT GTGGGAAAACAA

RG2B deduced polypeptide sequence (SEQ ID NO:90)

MSDPTGIAGAIINPIAQTALVPVTDHVGYMISCRKYVRVMQMKMTELNTSRISVEE
HISRNTRNHLQIPSQTKEWLDQVEGIRANVENFPIDVITCCSLRIRHKLGQKAFKITE
QIESLTRQLSLISWTDDPVPLGRVGSMNASTSASLSDDFPSREKTFTQALKALEPNQK
FHMVALCGMGGVGKTRMMQRLKKAAEEKKLFNYIVGAVIGEKTDPFAIQEAIADY
LGIQLNEKTKPARADKLREWFKKNSDGGKTKFLIVLDDVWQLVDLEDIGLSPFPNQ
GVDFKVLLTSRDSQVCTMMGVEANSIINVGLLTEAEAQSLFQQFVETSEPELQKIGE
DIVRKCCGLPIAIKTMACTLRNKRKDAWKDALSRIEHYDIHNVAPKVFETSYHNLQ
EEETKSTFLMCGLFPEDFDIPTEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQ
TNLLIESDDVGCVKMHDLVRAFVLGMFSEVEHASIVNHGNMPGWPDENDMIVHSC
KRISLTCKGMIEIPVDLKFPKLTILKLMHGDKSLRFPQDFYEGMEKLHVISYDKMKY
PLLPLAPRCSTNIRVLHLTECSLKMFDCSSIGNLSNLEVLSFANSHIEWLPSTVRNLK
KLRLLDLRFCDGLRIEQGVLKSFVKLEEFYIGDASGFIDDNCNEMAERSYNLSALEF
AFFNNKAEVKNMSFENLERFKISVGCSFDENINMSSHSYENMLQLVTNKGDVLDSK
LNGLFLKTEVLFLSVHGMNDLEDVEVKSTHPTQSSSFCNLKVLIISKCVELRYLFKL
NLANTLSRLEHLEVCECENMEELIHTGIGGCGEETITFPKLKFLSLSQLPKLSSLCHN
VNIIGLPHLVDLILKGIPGFTVIYPQNKLRTSSLLKEGVVIPKLETLQIDDMENLEEIW
PCELSGGEKVKLRAIKVSSCDKLVNLFPRNPMSLLHHLEELTVENCGSIESLFNIDLD
CVGAIGEEDNKSLLRSINVENLGKLREVWRIKGADNSHLINGFQAVESIKIEKCKRFR
NIFTPITANFYLVALLEIQIEGCGGNHESEEQIEILSEKETLQEATGSISNLVFPSCLMH
SFHNLRVLTLDNYEGVEVVFEIESESPTCRELVTTRNNQQQPIILPYLQDLYLRNMD
NTSHVWKCSNWNKFFTLPKQQSESPFHNLTTINILKCKSIKYLFSPLMAELLSNLKDI
RISECDGIKEVVSNRDDEDEEMTTFTSTHTTTTLFPSLDSLTLSFLENLKCIGGGGAK
DEGSNEISFNNTTATTAVLDQFELSEAGGVSWSLCQYAREIEIVGCYALSSVIPCYAA
GQMQKL

RG2C polynucleotide sequence (SEQ ID NO:91)

ATAATATTACACAAAGGTAACGTCATTAATTAATTACGATACGAGACAGA
CTTTTTCACTCGGACATNAACGGTCTATTCCTAACTTNANNTAATTNAAT
GAATTTAGGATGTGCTAATATGCATGTAANATTCGCTACCGTCATCTTTC
AAATGACCATATTTTTATGTATTTATAATGAATCAATGAAAAACCGGATT
TCTATTTAAAATTCTTAAAACTTCATCTTTTAAGCCAGGGTGAATACAAT
TGCTAGATCCACTGTTAATTTCCATCGAATTATGCCTGATCAATTGTTGG
CTGCCTACGATGCAGGTGCTACCACAAGAATATGGCCATGGAAACTGCTA
ATGAAATTATAAAACAAGTTGTTCCAGTTCTCATGGTTCCTATTAACGAT
TACCTACGCTACCTCGTTTCCTGCAGAAAGTACATCAGTGACATGGATTT
GAAAATGAAGGAATTAAAAGAAGCAAAAGACAATGTTGAAGAGCACAAGA
ATCATAACATTAGTAATCGTCTTGAGGTTCCAGCAGCTCAAGTCCAGAGC
TGGTTGGAAGATGTAGAAAAGATCAATGCAAAAGTGGAAACTGTTCCTAA
AGATGTCGGCTGTTGCTTCAATCTAAAGATTAGGTACAGGGCCGGAAGGG
ATGCCTTCAATATAATTGAGGAGATCGACAGTGTCATGAGACGACACTCT
CTGATCACTTGGACCGATCATCCCATTCCTTTGGGAAGAGTTGATTCCGT

GATGGCATCCACCTCTACGCTTTCAACTGAACACAATGACTTCCAGTCAA
GAGAGGTAAGGTTTAGTGAAGCACTCAAAGCACTTGAGGCCAACCACATG
ATAGCCTTATGTGGAATGGGGGGAGTGGGGAAGACCCACATGATGCAAAG
GCTGAAGAAGGTTGCCAAAGAAAAGAGGAAGTTTGGTTATATCATCGAGG
CGGTTATAGGGGAAATATCGGACCCCATTGCTATTCAGCAAGTTGTAGCA
GATTACCTATGCATAGAACTGAAAGAAAGCGATAAGAAAACAAGAGCTGA
GAAGCTTCGTCAAGGGTTCAAGGCCAAATCAGATGGAGGTAACACTAAGT
TCCTCATAATATTGGATGATGTCTGGCAGTCCGTTGATCTAGAAGATATT
GGTTTAAGCCCTTCTCCCAATCAAGGTGTCGACTTCAAGGTCTTGTTGAC
TTCACGAGACGAACATGTTTGCTCAGTGATGGGGGTTGAAGCTAATTCAA
TTATTAACGTGGGACTTCTAATTGAAGCAGAAGCACAAAGATTGTTCCAG
CAATTTGTAGAAACTTCTGAGCCCGAGCTCCACAAGATAGGAGAAGATAT
TGTTAGGAGGTGTTGCGGTCTACCCATTGCCATCAAAACCATGGCGTGTA
CTCTAAGAAATAAAAGAAAGGATGCATGGAAGGATGCACTTTCTCGTTTA
CAACACCATGACATTGGTAATGTTGCTACTGCAGTTTTTAGAACCAGCTA
TGAGAATCTCCCGGACAAGGAGACAAAATCTGTTTTTTGATGTGTGGTT
TGTTTCCCGAAGACTTCAATATTCCTACCGAGGAGTTGATGAGGTATGGA
TGGGGCTTAAAGTTATTTGATAGAGTTTATACAATTATAGAAGCAAGAAA
CAGGCTCAACACCTGCATTGACCGACTGGTGCAGACAAATTTACTAATTG
GAAGTGATAATGGTGTACATGTCAAGATGCATGATCTGGTCCGTGCTTTT
GTTTTGGGTATGTATTCTGAAGTCGAGCAAGCTTCAATTGTCAACCATGG
TAATATGCCTGGGTGGCCTGATGAAAATGATATGATCGTGCACTCTTGCA
AAAGAATTTCATTAACATGCAAGGGTATGATTGAGTTTCCAGTAGACCTC
AAGTTTCCTAAACTAACGATTTTGAAACTTATGCATGGAGATAAATCGCT
AAAGTTTCCTCAAGAATTTTATGAAGGAATGGAAAAGCTCCGGGTTATAT
CATACCATAAAATGAAGTACCCATTGCTTCCTTTGGCACCTCAATGCTCC
ACCAACATTCGGGTGCTTCATCTCACGGAATGTTCATTAAAGATGTTTGA
TTGCTCGTGTATTGGAAATCTATCGAATCTGGAAGTGCTGAGCTTTGCTA
ATTCTTGCATTGAGTGGTTACCTTCCACGGTCAGAAATTTAAAAAAGCTA
AGGTTACTTGATTTGAGATTGTGTTATGGTCTCCGTATAGAACAGGGTGT
CTTGAAAAGTTTGGTCAAACTTGAAGAATTTTATATTGGAAATGCATATG
GGTTTATAGATGATAACTGCAAGGAGATGGCAGAGCGTTCTTACAACCTT
TCTGCATTAGAATTCGCGTTCTTTAATAACAAGGCTGAAGTGAAAAATAT
GTCATTTGAGAATCTTGAACGATTTAAGATCTCAGTGGGATGCTCTTTTG
ATGGAAATATCAATATGAGTAGCCACTCATACGAAAACATGTTGCGATTG
GTGACCAACAAAGGTGATGTATTAGACTCTAAACTTAATGGGTTATTTTT
GAAAACAGAGGTGCTTTTTTTAAGTGTGCATGGCATGAATGATCTTGAAG
ATGTTGAGGTGAAGTCGACACATCCTACTCAGTCCTCTTCATTCTGCAAT
TTAAAAGTCCTTATTATTTCAAAGTGTGTAGAGTTGAGATACCTTTTCAA
ACTCAATGTTGCAAACACTTTGTCAAGACTTGAGCATCTAGAAGTTTGTA
AATGCAAGAATATGGAAGAACTCATACATACTGGGATTGGGGGTTGTGGA
GAAGAGACAATTACTTTCCCCAAGCTGAAGTTTTTATCTTTGAGTCAACT
ACCGAAGTTATCAGGTTTGTGCCATAATGTCAACATAATTGGGCTACCAC

ATCTCGTAGACTTGAAACTTAAGGGCATTCCAGGTTTCACAGTCATTTAT
CCGCAGAACAAGTTGCGAACATCTAGTTTGTTGAAGGAAGAGGTAGATAT
ATGTTCTTTATGTTAATACAATTTAAACAATATTTTCAACCAAATTTTCA
TAATATATCTGTAATTTGATTGTATGATGTGTTATTGTTTATATGTGGCT
ATTAAGGGATGATAATTTTGCAGGTTGTGATTCCTAAGTTGGAGACACTT
CAAATTGATGACATGGAGAACTTAGAAGAAATATGGCCTTGTGAACTTAG
TGGAGGTGAGAAAGTTAAGTTGAGAGAGATTAAAGTGAGTAGCTGTGATA
AGCTTGTGAATCTATTTCCGCGCAATCCCATGTCTCTGTTGCATCATCTT
GAAGAGCTTACAGTCGAGAATTGCGGTTCCATTGAGTCGTTATTCAACAT
TGACTTGGATTGTGTCGGTGCAATTGGAGAAGAAGACAACAAGAGCCTCT
TAAGAAGCATCAACGTGGAGAATTTAGGGAAGCTAAGAGAGGTGTGGAGG
ATAAAAGGTGCAGATAACTCTCATCTCATCAATGGTTTCAAGCTGTTGA
AAGCATAAAGATTGAAAAATGTAAGAGGTTTAGAAATATATTCACACCTA
TCACCGCCAATTTTTATCTGGTGGCACTTTTGGAGATTCAGATAGAAGGT
TGCGGAGGAAATCACGAATCAGAAGAGCAGGTAACGCTTTCAATTTCACT
TTCTTAATTAATTANGGACTAAGCTCCTGTTTTTGAATAATAAAGAGGT
GGGATGACTAAACTTGGGCATCACAATTGCAACAAAATGTTACAAACCAT
GAAACGCTCAAACCATTTCTTGAATTAAGGTTTCAATACAAGTCATTTAA
AAATATGGCTTAAATTTTTTTATATTTATGTATCAACATGATTTTTCATT
AGAGATCATTATTATAATAGTAAGTTTAAAGCAATTTAAATTAGAACTAA
TTCTAACTTTAGCTAATAAATCGTTATAAATGTAAATAATTACTTTTTAG
TGAAATAAGCAACGGATTTAATAAGTTAACAACTTAAATGTCATTTCCTA
ACAAAAAAAACTATTTGGTTCAGAAAAACTGTAATTCAAGATAACTAAAA
TAAAAATATTTGACATTCACTAAGAGCATTTTTTTCTAAATATGATTGCA
AATGAATAAAACTTAAATTTATACAGAAAAGATTTTTATATATGTTATAC
AAAATTTACAAATTGAAATTGGATATGTTAATTAACGGTTTATAATTCTG
GTATCACAAAGGGATATATAATAAAATATTATTTTTCTGTAGTCATTTAT
AATTGTACTAGTTTATAACCCGTGGGAACCATGAGTTCTAAAATTAGTTA
AACTTTCATAATAAAAATTTATAATTATTATTTATTTTAAATAAATTATT
AATTAAGAGATATATCAAAAATTTAAAGTTATTATAACTTCAAATTTAAC
ATATAATTAAAAAATATATGATCATAACTTTCCGCAACTCTTCTTTTGTA
TTAAAATGACCAGAGAAGCTCTTAGTATATTTCTAAATCAAAGTCACAA
AACTAATGAAGCATATAATTTTGTGAAAATCAATTAGCATTAGGTTTTAA
GAGTCACCAAATTCAAAGAGTAATCCAATGCTTTCATTACCACTATGGAG
AAAATATTTTCTTAGTTTAAATGAAATGAAAACAAACATTCAAACTAATT
GTTGCTTATTAAACCAAAGACCCATTACTTAGCCAAGAGTTTAACCAAAA
AAAATTACATTCATGTATCATTATTAATGACTAGATATATATGAATATGA
AGGGAGTTTTTATAGAAAATATAATCATAGATATTCAACATAACTTCATG
GAATTCCTCAAAATAACCAAGTTATTCAAGAAATTACATCCAAGTCAACC
AAAGAGAAGTTTAGCCTAGCATGGCTAAACTCAAGAAAATAAAATAAGGA
TTAGAAGTACCAAACATGTAGTAAGAATCACAGTAAAAGATGATGTTGTT
CTTGATGTTCTTCTAAGTTCTTCAAGTCTCCAGTTGCTCCTAATAATGCA
AAGGAGAGCCATTAAATTCGTATGTATTGATCCCTTCAAAAGCTGCACCA

ACCTCCCTTAAATAACACTCAAAGCAAAAATGACAAAATTGCCCCTGAAG
GACCCTATGCGGGTGCCTTGCGCGGGTGGAGCTGAATATGAAAGGTCTTT
GGTCTTTGTGAGGGTGATGTTGTGCGGGTTAGCTTGTCGCATGCTTCCGC
GCGGTTCGCGCACATGTGCACAGGTGATGCATGGTGTGTACGTTCTTGAC
TTTTGAGCCTCCGATGCTTAGTCCACTTGGCCCAATTCGAGTCCAATCAA
CTTATGACCCATTTTTCTTCAAGTTATCTTCAAGTTAAGCCCAATTTGCC
TTCTCCAAATCATCCATAACTTCACAGAATCGCCCGTTCATCTTAATCCC
GAATGAACAATTATTCTCCCGTCTTCATTTTAAGCAAGATACCACCTTCT
TCATGCTTCATCCATCAATAGTACACTTCATGTATCATCTCTACTAGTTA
TTTAGTCCACAGTCCTTGTTGTCCTCCAAATTTAATTATCTCATTTAGTT
CCCGTTCCGCTAGTTTCCTTAAAATTTGCAATTAAGCTCACAGAAATATT
AAGTACCCGAAATGGTCATAAAATAACAGAAAGGAAAATATGCATGAAGA
TTAACTAAATGATGAACGAAATATGCTAAAATAGACTATAAAATGAAGTA
AATAAAATGAAATTATCGCACTCCGACCACCCTTATAGGCTTGTAGTCCA
CCCACCCTTCATTCCTTGTACCAATATGGGATGGAAACATCATTAATTAA
GCCAAAAAACTAACATATAAGGGGTGAGTGACAAAGGTAAGTACTAAAGA
TGAAAAAAATCCATTTTTCTTGTATATACACAACACACACATAGGGGCAG
ACGTAGGATTTCATAGTACAGATTGTTGGTGGCACATAAGTGTTGCTAGT
GACATTTTTTTTTCTTTTACGTAGTGGCACAACAGTARAAAAAACRAAA
AATTCGAAATTTTTTACAATGTGCCTAAAAAAAACAGTGGTTGTTGGTGC
CACTATGGACACCAAAGTTGAACTGCCCCTGCGCGCGCACACACACACAC
ACATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTTTG
GGATGTGATACTTCTTTTGGGAAAATGGAGTTATATCTTTGATATTGTAT
TTTTTTAATGTAATTTATATATTTAATCATTTTAGTTTATAAGTTTTATT
TATTTKGATATGAAAAAAAAAGTCTTTTATACATTGGATTTAACATAAAA
ATCCAACAATATTAATCAAAAAGACCAAACATGTGGACAATTATGTATAT
AATTAATTCACAATAGTCTTTAGGAATAGNATTATATATATAATTAATTC
TCAATGGTCTTAGGAATAGTAAGTTCTTATATTTCAAACNTTTGCCACAN
TTCTTTGNTTACTTNGACACTTTYCTCTMWNNANWMWWTWATATATATAT
ATATATATATATAHAHAHAHAVACACACACACTAGATGTGTGCCMGCGCA
AAGCAGTGACGTNNNGGAGAANACTTTCTTAAGCATAAATAATTATTATA
TTTTTTATTGGGTATTATATAATAAAAAATTACAACTTTTAAATAAAATA
TTTATGTTTATACTTTATATTTATATTGCTTGTATACTATTAATATAATA
AATTAATATTTATGTCTAATTTATGAAATGTAAATTAATTTAAATACATG
AATTTAATATTTTTAAAATTTTCAGTTTGCTTCAAATTGAGTTTCTTAAT
TATTGACCAAACATGTGGACAATTATGTATATAATTAATTCACAATAGTC
TTTAGGAATAGTATTATATATATAATTAATTCTCAATGGTCTTAGGAATA
GTAAGTTCTTATATTTCAAACTTTTGCCACAATTCTTTGCTTACTTTGAC
ACTTTTCCTTCCTAACTTTACATATATATATATATTAAAGCGCAAAGGTC
ATAGGAATATAATATTTTCTATTATTCTACGTTTTGCCACAAAAGTTTGA
ACACTTTGCCACTTTTTGTCCCTCCTTAACCTTTTCAATGTTTTGCGACA
AAAGTTCCAAAACTTTGCCACTTTGATCATTCCTCAACTTTTCACCGCAT
TAGTTTGTGGAGTTGGCAGTTTTGGTCCCTCTAACTTCGATATTCTCTAC

TGCTAGCCAAAAAGGGTTCCAGAGTTTCACACTTTTGGTCCCTGACAGTA
ACCAAATGTGAGATGTCAAATTTTTGCCACATTAGTTTGTGGAGTTGTCC
CTTTTGGTCCCCCCACATTCGATATTCTACTATACGATCTTATTTTTCTC
AAATAACAACACGTATATTTAATTACTAATGATAGAAATAGATATCAAAT
AAAGTATTTGTAACACTGTGTAGAGTTTTTTTTACAAGTTTGTATCAAA
TCATATCAAAATTTAAGGTGGAACGGTGACCACATTAACCAGAAATGTAA
TTTATTCTTTGATTTTGATAATTTTTAATATTTTGTTGTGATCTATGTAT
TTAAAAGTAAACAACAAAGAACATAATCCAAAACCCTAAATTGCAAGTCT
CGCCCAATTTCTCTATCACTAGTCCTCACTTACGATGGCGTTACGTCGCT
CTCTCACTGCTTACAACCCTTTGTTGCTACTCATTACAATAACGAAAAGT
TGAATATCCATATATTTATTTGGATGTGGAATTGAACGAATCTCGTCAAA
TTTTTGATTTAGTTGATGGATTTGAGTAGAAGTTTGGGCAGAACGGGAAT
GATGGTCTGCAAGTGGTTATAAACTTGATTCTGAGTTATTACTATATATG
TAGCCTCTTTACAACGACCAAGGTTTCTTCCAGGTACCATTTGATCTTTT
TAGAACTTAGTTTTCTGAAACACCCTGATTTGGATCAAATATCACCAACA
ACTCTTAAAAACTTGATTAATCAATTGTTTACTTCATCTTGATAACAAGT
GGAATGATTTCTACTTGAAAAAAAAGGTCCATGTGCGTCTGGTGGATCT
GGTAAATGAAGATGGAAGGGAGAGCTGACTTTAAAGACACAAACACGTCA
CCATATCTTTTATTTTATTTTAAATTTTCTTTTTTCCTATTTCTTTCTTT
CTTGATCTCCAGATGGTATGTGGTGTGGATAATTTACACATAGAGATTGG
GAACGACTGTGATTTAGAGAGGACGTGGCTTGGGGTTGAGGATGGTTTAT
GGCTGGCCGAGTTTCATTTATATAAACAAACAAATATATAAAACAAGGGG
TAAAATGGCCATCTTATATGTATTTAACCGTCCTCTTTTTTTTTTTTTTT
TTTTTTTTTTTTTTTTTTGTAATTTAAGAAGGGGTATACCAGTGTCAGC
CTCTTATTCCCAACCAGTCAAATAGGGACTTAGGTTGTTTGGAAACAGTT
CCGTGAGACCGTGACTTGGATGGTAGATAAATTTAGTAAACTTAACCCTT
CAATTAACCTACCTTTTTCTTATTAACTCAATTTCAACCTAAATTCTGAT
TCTTGTTTGAAAATAAGTTGCATCTTTATTTTTGTATTATCTTGTTGCAT
AGGATCCTTAGCATCTTTTAATAATTTATTTGAAGGTGAAAGATCCAACT
ATTTTTAATCTGTTGACGTTTTCCATCATTTGCAACTGTTTCTTGAAAAA
AAAATACCTAAAATCAAAATAACCATTTTCAAATCCAAAATTATAAGAGA
GAATTGTAAATGGACATGGAATCATAAATCATTAACACAGTTCAGTAAAC
AAGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTATCAGAGAA
AGAGACATTACAAGAAGCCACTGGCAGTATTTCAAATCTTGTATTCCCAT
CCTGTCTCATGCACTCTTTTCATAACCTCCGTGTGCTTACATTGGATAAT
TATGAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTGAGAGTCCAACAAG
TAGAGAATTGGTAACAACTCACAATAACCAACAACAGCCTATTATACTTC
CCTACCTCCAGGAATTGTATCTAAGGAATATGGACAACACGAGTCATGTG
TGGAAGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATC
AGAATCACCATTCCACAACCTCACAACCATAGAAATGAGATGGTGTCATG
GCTTTAGGTACTTGTTTTCGCCTCTCATGGCAGAACTTCTTTCCAACCTA
AAGAAAGTCAAGATACTTGGGTGTGATGGTATTAAAGAAGTTGTTTCAAA
CAGAGATGATGAGGATGAAGAAATGACTACATTTACATCTACCCACAAAA

CCACCAACTTGTTCCCTCATCTTGATTCTCTCACTCTAAACCAACTGAAG
AATCTGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGGGGAGCAATGA
AATATCTTTCAATAATACCACTGCAACGACTGCTGTTCTTGATCAATTTG
AGGTATGCTTTGTACATATTCAATTATTTATTTAATTTCCTTTTTTATTT
GCAATATTCTATAAATAATACATTTTATACCCACTATACTAAGATAATAA
TTACCTAGAGGGATGGATGCTATGACACAGCTGCTACACTTCAGAAACTC
TAGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTTTTGA
TGGGTAATATAGGCAATTTAAGTTTTATTTCTGTTAAAGCAGTATTTAGC
AAGTACTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTGAAAATCTGGT
CATTGTACCCAGAATTTAGTTAAATGTAACATTTTAGATATTAGGGGTTA
TCAGGTGACAGATATTGTAGAATAGAACAATATGTAATATTACCCAAAAC
TATTTTTTCTAAGGTTGCTCTGTTAAATATGTGCTTTCTTGATTTCATTG
AATTTGCATTCCTATATTTTAGGTGGTAAAGTGATTGTCTCTTCAATAAA
TCCCGAAATTAATTAAAAAAAAAAAAACAAAAGTAAATTTTTGATATGGA
GAGCACTGGTATCATTTAGTATATAAAAAAACTAGATTTTGAATTAAGTT
TCTTATATAAAAGCTGTGTATATAGTTTAATTAGTTTTACATCATTTTTC
CATGTGGTGTTGCAGTTGTCTGAAGCAGGTGGTGTTTCTTGGAGCTTATG
CCAATACGCTAGAGAGATAAAAATAGGCAACTGCCATGCATTGTCAAGTG
TGATTCCATGTTATGCAGCAGGACAAATGCAAAAGCTTCAAGTGCTGAGA
GTAATGGCTTGCAATGGGATGAAGGAGGTATTTGAAACTCAATTAGGGAC
GAGCAGCAACAAAAACAACGAGAAGAGTGGTTGTGAGGAAGGAATTCCAA
GAGTAAATAACAATGTTATTATGCTTCCCAATCTAAAGATATTAAGTATT
GGAAATTGTGGGGGTTTGGAACATATATTCACATTCTCTGCACTTGAAAG
CCTGAGACAGCTCCAAGAGTTAACGATTAAGGGTTGCTACAGAATGAAAG
TGATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGCAAACAACAACA
ACAACAACGAAGGGGGCATCTTCTTCTTCTTCTTCTAAGAAGGTGGT
GGTCTTTCCTTGTCTAAAGTCCATTGTATTGGTCAATCTACCAGAGCTGG
TAGGATTCTTCTTGGGGATGAATGAGTTCCGGTTGCCTTCATTAGATAAA
CTTATCATCGAGAAATGCCCAAAAATGATGGTGTTTACAGCTGGTGGGTC
CACAGCTCCCCAACTCAAGTATATACACACAAGATTAGGCAAACATACTC
TTGATCAAGAATCTGGCCTTAACTTTCATCAGGTACATATATATTCCTTT
AATTGGCATCATCTAATTAAGAAAGATATCATTCCTGCCAAGTAAATTTA
CTTCAAACACATTCACACTAGTTTCAGTCCAAGTTTATGTTGTTCTAGGA
AGGCCAAAATGGGAAAGCAAGATAGGGAAAAATAGAGTATTTCAGTGGAA
AGGGTATTTTAGGTATTTTCTGTCAAAAATTGTTATTGCAGGCTTTTTAG
TACCTGGAAGAGCATGATTATTCTCGATTTGCTTGTTTCTTTATCATTTT
TCTTAGCCTAGCATGATTTTCAATGAAATCTTTCCCTGTTACTCCATTTG
ATTGTTGTTCTTCATGGTTCTAAGTGAGTTAGTGGCTCATCTGTTACTTC
TTTTGATTGTTATTTTCATAGCATGTTGTCACTTGAATCAAGCTTTTCCA
TTTTCAACAAGGACAAAAGGTCAAAACTAACCTACTTTATGAGATCAAAA
ATAGCAACCCATCGGATAACTTTTAGTTGGAGTTAATACTTACAATTACC
ATTGTGATTAATAATTATAATATCTTGTATTAATTCATAAAAATTGGTAC
AGCACATATATGACATTTCAAAGGTTTTTGTTTGACATATATATGCCTCT

GGCGTTTTCTTTATTGGACATGCAGACTTCATTCCAAAGTTTATACGGTG
ACACCTTGGGCCCTGCTACTTCAGAAGGGACAACTTGGTCTTTTCATAAC
TTTATCGAATTAGATGTGAAATCTAATCATGATGTTAAAAAGATTATTCC
ATCCAGTGAGTTGCTGCAACTGCAAAAGCTGGTAAAGATTAATGTAATGT
GGTGTAAAAGGGTAGAGGAGGTATTTGAAACTGCATTGGAAGCAGCAGGG
AGAAATGGAAATAGTGGAATTGGTTTTGATGAATCGTCACAAACAACTAC
CACTACTCTTGTCAATCTTCCAAACCTTGGAGAAATGAAGTTACGGGGTC
TCGATTGTCTGAGGTATATATGGAAGAGCAATCAGTGGACAGCATTTGAG
TTTCCAAACCTAACAAGAGTTGAAATTTATGAATGTAATTCATTAGAACA
TGTATTTACTAGTTCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTAG
AGATTGGTTTGTGCAACCATATGGAGGTCGTGCATGTTCAGGATGCAGAT
GTTTCTGTAGAAGAAGACAAAGAGAAAGAATCTGATGGCAAGATGAATAA
GGAGATACTTGTGTTACCTCATCTAAAGTCATTGAAATTACTACTTCTTC
AAAGTCTTAAGGGGTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATTA
TTGGATACTTTGGAAATCTACGAATGCCCAGCAATAACCACCTTCACCAA
GGGAAATTCCGCTACTCCACAGCTAAAAGAAATGGAAACAAATTTTGGCT
TCTTTTATGCTGCAGGGGAAAAAGACATCAACTCCTCTATTATAAAGATC
AAACAACAGGTAAACCAGATCTTTGTTGCTTTAATAATTCTTAAACTACA
TTTGAAAAGCTTCATGCAAGTTTTTTTGTTATATTGTCAAAAACCGCAA
CCTACATTTTCAGCTTTATATTTATGTACTTTATGCAGGATTTCAAACAA
GACTCTGATTAATGTGAAGTGAATATTAAAGGTAAATTATATTTTCATGT
TCCTAGTTGCCTATTAATTAAAGGCCTTTTAGTTCGTGATTTTTGGATGT
ATTCTTCATGATGATGTCAATCTTCTAATACCCCATTCATTGTTTGGTTG
AATGTTGACTCTATGTCAGGATGAATATTCAAGGGAAGAATTGTTCATCA
TATGAAGGACATTAAAGAACATGGTGCTAT

RG2C deduced polypeptide sequence (SEQ ID NO:92)

MAMETANEIIKQVVPVLMVPINDYLRYLVSCRKYISDMDLKMKELKEAKDNVEEH
KNHNISNRLEVPAAQVQSWLEDVEKINAKVETVPKDVGCCFNLKIRYRAGRDAFNI
IEEIDSVMRRHSLITWTDHPIPLGRVDSVMASTSTLSTEHNDFQSREVRFSEALKALE
ANHMIALCGMGGVGKTHMMQRLKKVAKEKRKFGYIIEAVIGEISDPIAIQQVVADY
LCIELKESDKKTRAEKLRQGFKAKSDGGNTKFLIILDDVWQSVDLEDIGLSPSPNQG
VDFKVLLTSRDEHVCSVMGVEANSIINVGLLIEAEAQRLFQQFVETSEPELHKIGEDI
VRRCCGLPIAIKTMACTLRNKRKDAWKDALSRLQHHDIGNVATAVFRTSYENLPD
KETKSVFLMCGLFPEDFNIPTEELMRYGWGLKLFDRVYTIIEARNRLNTCIDRLVQT
NLLIGSDNGVHVKMHDLVRAFVLGMYSEVEQASIVNHGNMPGWPDENDMIVHSC
KRISLTCKGMIEFPVDLKFPKLTILKLMHGDKSLKFPQEFYEGMEKLRVISYHKMKY
PLLPLAPQCSTNIRVLHLTECSLKMFDCSCIGNLSNLEVLSFANSCIEWLPSTVRNLK
KLRLLDLRLCYGLRIEQGVLKSLVKLEEFYIGNAYGFIDDNCKEMAERSYNLSALEF
AFFNNKAEVKNMSFENLERFKISVGCSFDGNINMSSHSYENMLRLVTNKGDVLDSK
LNGLFLKTEVLFLSVHGMNDLEDVEVKSTHPTQSSSFCNLKVLIISKCVELRYLFKL
NVANTLSRLEHLEVCKCKNMEELIHTGIGGCGEETITFPKLKFLSLSQLPKLSGLCH

NVNIIGLPHLVDLKLKGIPGFTVIYPQNKLRTSSLLKEEVVIPKLETLQIDDMENLEEI.
WPCELSGGEKVKLREIKVSSCDKLVNLFPRNPMSLLHHLEELTVENCGSIESLFNID
LDCVGAIGEEDNKSLLRSINVENLGKLREVWRIKGADNSHLINGFQAVESIKIEKCK
RFRNIFTPITANFYLVALLEIQIEGCGGNHESEEQIEILSEKETLQEATGSISNLVFPSC
LMHSFHNLRVLTLDNYEGVEVVFEIESESPTSRELVTTHNNQQQPIILPYLQELYLR
NMDNTSHVWKCSNWNKFFTLPKQQSESPFHNLTTIEMRWCHGFRYLFSPLMAELL
SNLKKVKILGCDGIKEVVSNRDDEDEEMTTFTSTHKTTNLFPHLDSLTLNQLKNLK
CIGGGGAKDEGSNEISFNNTTATTAVLDQFELSEAGGVSWSLCQYAREIKIGNCHAL
SSVIPCYAAGQMQKLQVLRVMACNGMKEVFETQLGTSSNKNNEKSGCEEGIPRVN
NNVIMLPNLKILSIGNCGGLEHIFTFSALESLRQLQELTIKGCYRMKVIVKKEEDEYG
EQQTTTTTTKGASSSSSSSKKVVVFPCLKSIVLVNLPELVGFFLGMNEFRLPSLDKLII
EKCPKMMVFTAGGSTAPQLKYIHTRLGKHTLDQESGLNFHQTSFQSLYGDTLGPAT
SEGTTWSFHNFIELDVKSNHDVKKIIPSSELLQLQKLVKINVMWCKRVEEVFETALE
AAGRNGNSGIGFDESSQTTTTTLVNLPNLGEMKLRGLDCLRYIWKSNQWTAFEFPN
LTRVEIYECNSLEHVFTSSMVGSLLQLQELEIGLCNHMEVVHVQDADVSVEEDKEK
ESDGKMNKEILVLPHLKSLKLLLLQSLKGFSLGKEDFSFPLLDTLEIYECPAITTFTK
GNSATPQLKEMETNFGFFYAAGEKDINSSIIKIKQQDFKQDSD.

RG2D polynucleotide sequence (SEQ ID NO:93) and (SEQ ID NO:94)

ACGACCACTATAGGGCGAATTGGGCCCGACGTCGCATGCTCCCGGCCGCC
ATGGCCGCGGGATGTAAAACGACGGCCAGTCGAATCGTAACCGTTCGTAC
GAGAATCGCTGTCCTCTCCTTCAACCATTTAATGTATATGAGCTAAATTG
AAACATCTACTATCATGTTTAAATTTATAAACTTTTTCCTTTAGATTCAC
TTGTCTGGATGTGTTTAATAAAACCCAATTTCCCACATGCGTAGAGATCA
TAGATGTAACTATTGTTAATCAATTTTGCCTGCCAAGTTTTAATAATTAT
ACTTGGATATTAACAAAACTTTATCTAACGACCAAGGTAATATTAAAAAT
AGGTTATTATTCTTCATGCTAATTAAAAGATGGGTTGCAAAAGTGAGACC
ATGAAAACATTAACACGTTGATATTTTCAACTTTTATTCTTTCATATTCA
CCATATTTTTTACTTTCGTATTGATTAATCATCTTTCAATCACAGGCTCC
TTGGCAAAAAGTCAGATCTATTAACAAATACTTCCATGTGGTTGCAAATT
ACAAGGATTTCAACATAATTACCAAAACATAGCATTATCATAAGATCGAA
TAATAATCAAATTCTTCTATAATATTACACAAAGGTAACGTCATTAATTA
ATTACGATACGAGACAGACTTTTTCACTCGTGACATCAACGGTCTATTCT
AACTTTACTTAATTAAATGAATCTAGGATGTGCTCATATGCATGTAATAT
TTGCTACCGTCATCTTTCAAATGACCATATTTTTATGTATTTATAATGAA
TCAATGAAAAACCGGATTTCTATTTAAAATTCTTAAAACTTCATCTTTTA
AGCCAGGGTGAATACAATTGTAGATCCACTGTTAATTTCCATCGATTATG
CGTGATCAATTGTTGGCTGCATACGATGCAGGTGCTACCACAAGAATATG
GCCATGGAAACTGCTAATGAAATTATAAAACAAGTTGTTCCAGTTCTCAT
GGTTCCTATTAACGATTACCTACGCTACGTCGTTTCCTGCAGAAAGTACA
TCAGTGACATGGATTTGAAAATGAAGGAATTAAAAGAAGCAAAAGACAAT
GTTGAAGAGCACAAGAATCATAACATTAGTAATCGTCTTGAGGTTCCAGC

AGCTCAAGTCCAGAGCTGGTTGGAAGATGTAGAAAAGATCAATGCAAAAG
TGGAAACTGTTCCTAAAGATGTCGGCTGTTGCTTCAATCTAAAGATTAGG
TACAGGGCCGGAAGGGATGCCTTCAATATAATTGAGGAGATCGACAGTGT
CATGAGACGACACTCTCTGATCACTTGGACCGATCATCCCATTCCTTTGG
GAAGAGTTGATTCCGTGATGGCATCCACCTCTACGCTTTCAACTGAACAC
AATGACTTCCAGTCAAGAGAGGTAAGGTTTAGTGAAGCACTCAAAGCACT
TGAGGCCAACCACATGATAGCATTATGTGGAATGGGGAGAGTGGGGAAGA
CCCACATGATGCAAAGGCTGAAGAAGGTTGCCAAAGAAAAGAGGAAGTTT
GGTTATATCATCGAGGCAGTTATAGGGGAAATATCGGACCCCATTGCTAT
TCAGCAAGTTGTAGCAGATTACCTATGCATAGAGCTGAAAGAAAGCGATA
AGAAAACAAGAGCTGAGAAGCTTCGTCAAGGGTTCAAGGCCAAATCAGAT
GGAGGTAACACTAAGTTCCTCATAATATTGGATGATGTCTGGCAGTCCGT
TGATCTAGAAGATATTGGTTTAAGCCCTTCTCCCAATCAAGGTGTCGACT
TCAAGGTCTTGTTGACTTCACGAGACGAACATGTTTGCTCAGTGATGGGG
GTTGAAGCTAATTCAATTATTAACGTGGGACTTCTAATTGAAGCAGAAGC
ACAAAGATTGTTCCAGCAATTTGTAGAAACTTCTGAGCCCGAGCTCCACA
AGATAGGAGAAGATATTGTTAGGAGGTGTTGCGGTCTACCCATTGCCATC
AAAACCATGGCGTGTACTCTAAGAAATAAAAGAAAGGATGCATGGAAGGA
TGCACTTTCTCGTTTACAACACCATGACATTGGTAATGTTGCTACTGCAG
TTTTTAGAACCAGCTATGAGAATCTCCCGGACAAGGAGACAAAATCTGTT
TTTTTGATGTGTGGTTTGTTTCCCGAAGACTTCAATATTCCTACCGAGGA
GTTGATGAGGTATGGATGGGGCTTAAAGTTATTTGATAGAGTTTATACAA
TTATAGAAGCAAGAAACAGGCTCAACACCTGCATTGAGCGACTGGTGCAG
GCAAATTTACTAATTGGAAGTGATAATGGTGTACACGTCAAGATGCATGA
TCTGGTCCGTGCTTTTGTTTTGGGTATGTATTCTGAAGTCGAGCAAGCTT
CAATTGTCAACCATGGTAATATGCCTGGGTGGCCTGATGAAAATGATATG
ATCGTGCACTCTTGCAAAAGAATTTCATTAACATGCAAGGGTATGATTGA
GATTCCAGTAGACCTCAAGTTTCCTAAACTAACGATTTTGAAACTTATGC
ATGGAGATAAGTCTCTAAAGTTTCCTCAAGAATTTTATGAAGGAATGGAA
AAGCTCCAGGTTATATCATACGATAAAATGAAGTACCCATTGCTTCCTTT
GGCACCTCAATGCTCCACCAACATTCGGGTGCTTCATCTCACTGAATGTT
CATTAAAGATGTTTGATTGCTCTTCTATCGGAAATCTATCGAATCTGGAA
GTGCTGAGCTTTGCTAATTCTCGCATTGAATGGTTACCTTCCACAGTCAG
AAATTTAAAGAAGCTAAGGTTACTTGATCTGAGATTTTGTGATGGTCTCC
GTATAGAACAGGGTGTCTTGAAAAGTTTGGTCAAACTTGAAGAATTTTAT
ATTGGAAATGCATATGGGTTTATAGATGATAACTGCAAGGACATGGCAGA
GCGTTCTTACAACCTTTCTGCATTAGAATTCGCGTTCTTTAATAACAAGG
CTGAAGTGAAAAATATGTCATTTGAGAATCTTGAACGATTCAAGATCTCA
GTGGGGTGCTCTTTTGATGGAAATATCAGTATGAGTAGCCACTCATACGA
AAACATGTTGCAATTGGTGACCAACAAAGGTGATGTATTAGACTCTAAAC
TTAATGGGTTATTTTTGAAAACAGAGGTGCTTTTTTTAAGTGTGCATGGC
ATGAATGATCTTGAAGATGTTGAGGTGAAGTCGACACATCCTACTCAGTC
CTCTTCATTCTGCAATTTAAAAGTCCGTATTATTTCAAAGTGTGTAGAGT

TGAGATACCTTTTCAAACTCCATGTTGCAAACACTTTGTCAAGCCTTGAG
CATCTAGAAGTTTGTGGATGCGAAAATATGGAAGAACTCATACATACTGG
GATTGGGGGTTGTGGAGAAGAGACAATTACTTTCCCCAAGCTGAAGTCTT
TATCTTTGAGTCAACTACCGAAGTTATCAGGTTTGTGCCATAATGTCAAC
ATAATTGGGCTACCACATCTCGTAGACTTGAAACTTAAGGGCATTCCAGG
TTTCACAGTCATTTATCCGCAGAACAAGTTGCGAACATCTAGTTTGTTGA
AGGAAGAGGTAGATATATGTTCTTTATGTTAATACAATTTAAATAATATT
TTCAACCAAAATTTCATAATATATCTGTAATTTGATTGTATGATGTGTTA
TTGTTTATATGTGGCTATTAAGGGATGATTATTTTGCAGGTTGTGATTCC
TAAGTTGGAGACACTTCAAATTGATGGCATGGAGAACTTAGAAGAAATAT
GGCCTTGTGAGCTTAGTGGAGGTGAGAAAGTTAAGTTGAGAGAGATTAAA
GTGAGTAGCTGTGATAAGCTTGTGAATCTATTTCCGCACAATCCCATGTC
TCTGTTGCATCATCTTGAAGAGCTTAAAGTCAAAAATTGTCGTTCCATTG
AGTCGTTATTCAACATCGACTTGGATTGTGTCAGTGCAATTGGAGAAGAA
GACAACAAGAGCATCTTAAGAAGAATCAAAGTGAAGAATTTAGGGAAGCT
AAGAGAGGTGTGGAGGATAAAAGGTGCAGATAACTCTCGTCCCCTCATCC
ATGGCTTTCCAGCTGTTGAAAGCATAAGTATCTGGGGATGTAAGCGGTTT
AGAAATATATTCACACCTATCACCGCCAATTTTGATCTGGTGGCACTTTT
GGAGATTCACATAGGAAATTACAGAGAAAATCATGAATCGGAAGAGCAGG
TAACGCTTTCAATTTCACTTTCTTACTTAATTAAGGACTAAGCTCTTGTT
TTTTGAATAATAAAGAGGTGGGATGACTAAACTTGGGCATCACAATTGTA
ACAAAATGTTACAAACCATGAACGTACAAACCATTTCTTGAATTAAGGTT
TCAATACAAGTCATTTACAAATATGGCTTAAGTTTTTTTATATTTATGTA
TCAACATTATTTTTCATTAGAGGTCATTATTATAATAGTAAGTTTAAAGC
AATTTAAATTAGCACTAATTTTTCATCATCTAACTTTAGCTAATAAATCG
TTATAAATGTCAATAGCTAAAATAAAAATATTTGACATTCACTGAGAGCA
ATTTTTTCTAAACATGATTGCAAATGATTAAAACTTAAATTTAAACTAAA
AAGATTTTTATATATGTTATACAAAATTTACAAATTGAAATTGGATATGT
TAATTAACAGTTTATAATTATTGTATTACAAAGCGATATATAATAAAATA
TTATTTTTCTGTAGTCATGTATAATTGTATATGTAAATGATTTTTTAAGA
TGGTAGAAGTGGAAACTAGTCAATCTCACTTAACTCATTGTCACACCAGT
TTTATATCCGTTTCTCTCTCTCTCTCTTCTTGCCTCCATCTTTTTTCAAC
TCATAACACATAAAAATAACATATTTTCCAACACATTTAAGTCACTACCA
CATCATTATTTTTAATTTAATTAAATTAGAAAATATAAAATTAAATAAAA
CATAACATTTTTTTATTAAAAGGCACTAATACAAATAAAAAGATACACGG
TAAATAAAAAAACGATAATTAGAAAAAAAACATAATAAAAAAAGACAACA
TTAAAAATAWAAAGCGACAACTAAAATTAACTAATGATCAAGAAAATTCT
AAAACTCCCACCATATTTTTCTGCAATTTGTCATTTATGTTCAAACACCA
TTCGCAGAATCCCTCCTATCAAGTGATCATGTTGATTGAGAAAAAACTGT
ATGTCTCTCTCATGTATCTCCAAGTCCAACAAGTTAGCTTTCATTTCTTC
ATTTTCTCATGTAAGACGCAAATTTTCATCCCGATATTGTTTTCTATCTT
CCACCTCTACTTTATTCACAGTGTGGATGAAGGAGAGGACAGCGATTCTC
GTACGAACGGTTACGATTCGACTGGCCGTCGTTTTACAATCCCGCGGCCA

TGGCGGCCGGGAGCATGCGACGTCGGGCCCATTCGCCCTATAGTGGTCGT
AATACA **(SEQ ID NO:93)**

Sequence gap

TGAGCCTCCGATGCTTAGTCCACTTGGCACAGTTCAAGTCCAATCAACTT
ATAACCCATTTTTCTTCAAGTTGTCTTCAAGTTAAGCCCAATTTGCCTTC
TCCAAATCATCCATAACTTCATGGAATCGCCCCTTCATCTTAATCCCGAA
TGCACAATTATTCTCCCATCTTCATTTAAGCAAGAGGCCACCTTCTTCA
TGCTTCATCCATCAATAGTCTGTTGGAATAGTGTCTAAGGCTGCAACTAT
ATTAGACAAGTATTTGACCCGGTTGTGCATGGTCCTTTTGGGTTGCCTTC
ACCATAGCAACTTGATAGGATGATTTATTAAGAGAGAGTAAATATTATTA
ATATATTATGAGAATAATATAATGAATAATATATTTGTTATTTGATTAAT
ATAAGTCATAGAATTAATTAGAATTAATTTGGTGACTTAAAGAGATTAAT
TAAATAAAGGGGTATAAACTGTCAATTGTTTGATAGTTAAGCTTTAGACT
GTAAATCCATTTGGATATGGTATGGACGAATCCTAAGGGATTTAGGATAG
CTAAAATCGTCCATATGAGTTATCTAAGAAGGATTTGGATAGCCTTAAGA
GAAGATTATCTGATAGGGACTTATCTGTAATCCTTAAGGAGTCTACAAGT
ATAAATAGACCCTATGGCTGATGGAATTCGACACATCTCCTAAAGTAAGA
GAGCCTTGGCCGAATTCCTCCCCTCACCTCTCTCCTAAATCATTCTTCTT
GCTATTGGTGTTTGTAAGCCATTAGAGGAGTGACATTTGTGACTCTAGAA
TCTCCAAGACCTCAAGATCAACAAGGAATTCAAAGGTATGATTCTAGATC
TGTTTCAATGTTGTTATTTGTCCTAATTAGTCATTAGAAGACTTGGATTC
AAAGCATGTTTATTAGAAAGCCTAGATCYGAGCAATAGGGTTTTGCATGC
GCACATAGGAAAGTTCTTATGGCTAAAACCCATCATAGTCCACTTCATGT
ATCATCTCTACTAGTTATTTAGTCCATAATCCTTGTTGTCCTCCAAGTTT
AATTACCTCCCTTAGTTCCTGTTCTGCTAGTTTCCTTAAAATTTGCTATT
AAGATCACAGAACTAGAGAGTACCCAAAATGGTTATAAAATAACAAAAAG
GAAAATATGCATGAAGATTAACTAAATTATAAATGTAATATGCTAAAATA
AACTATAAAAAAAAAGTAAATAAAATGAAACTATCACACTCCGACCACCC
TTATAGGCTTGTACTGCACCCACCCTTCATTCCTTGTACCAATATGGGAT
GGAAACATTATTCATTAAGCCAAAAAACTAACATTTAAGGGGTGAGTGAC
AAAGGTAAGTACTAAAGACAACAATAATCCATTTTTCTTGTACATACACA
ACACACACATAGGGGCGGACGTAGGATTTGTAGTATGTGTTGTGGGTGAC
ACATTTTTTCTTTTACGTAGTGACACAATAGTAGAGAAAACGAGAAATTC
CAATTTTTTACATTGTGTTCGAAAAAATATACAGGGGTTGCTGGTGCTAC
TCTGGGCACCAAAGTGGAACCGCCCCTGCACACACACACACATAGAGGGA
GAGAGAGAGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTT
TGGGATGTGATACTTCTTTTGGGAAAATGGAGCAATATCTTTAATATTGT
ATTTTTTTAATGTAATTTATATATTTAATCATTTAGTTTATAACTTTTA
GTTTTTTTTATTTTAATCTGTATATTTAATCATTTCAGTTTATAAGTTTT
ATTTATTTTGGTATACCAGAAAAAAAAGTCTTTTATGTGTTGGATTTAAC
ATAAAAATCTAACAATATTAATCAAAAAGACCAAACATGTGGACAATTAT
GTATATAATTAATTCTCAATGGTCTTAGTGTAACGATATAAATTTCAAAA
CAATTTTTCACATTAAAAAAAACACTTTCAGTCATAATTGTTATAAATTA

TCATTGTATCACAAAATCAGTTCATAACATCACATCCCAAGATCAATAAA
GTGTAAATACTCCTCATGTGTGTACTAATCAAGCCGACGCCTTCCCGCGA
TTCTCACTGGTACCTGAAACACGTAACATAACAACTGTAAGCATAAATGC
TTAGTGAGTTCCCCAAAATACCACATACCACATATATGCCTTTCCAGGCC
ATAACTCTGTAGGATCTTCCGACCCAAGTGTCTCAGGGGACTTCCGTCCC
GAATCCCGGTAGACCTTCCGGTCCTACCCGTATTGACCTTCCGGTCCGTA
TCATACATAACATACATAACACATACATATCACATAACAACATATAGCAC
ATACATCTCATAACATAAAAGACCTTCCGGTCACATAAAGGTACCCTTCC
AGGTACAGTATAGTGAGAANACTCACCTCGTATGATGTCTAATACCTCAC
GTGCTCGATATCCCTGAATCTCGAAACAATGACCTAGCCCCGCCTACTCA
CATAAAGTAATTATTTCAAATCATTAACGGCTCTCAAGGCTAGACTACAT
CCCTTTCTATAAATCCACAGAAGGGTAAAAGACCATTTTACCCCTCCTTG
ACCCAAAAGTCCAAATGTTGATCAAAACCCCAAAAGTCAACGAAAGACAA
TGGTCAACTTTGACCCTACTCGTGGAGTGCACAAAGGTGACTCGGCAAGT
ACATGCGGGTCCTCTGAATCCTTTCAGTCTCTCTTGGCTCGTCGAGTCTT
TCTTCCACCCGACGAGTTACACCTGTCATGAATCGCGGGGCAACCCCGAC
TCGACTTGTCGAGTCCGCTCATGGACTCAACGAGTTCATTCCATGCTCAC
ACTCAAATGACCTCCTGAGGTCAGATCTGTTCCTCTAATCCATAGATCTG
ACCTTCCCAAGCTCAATAAACACGTAAAGGTTCGAACTTGATACTCATGC
AACGTCCAAATGATTCTACTTGATGATTTAGCCCCAAATACAACATCCTA
AGTCCATACGACCTTATTTTTCTCAAATAACAACACATATATTTAATTAC
CAATGACAGTAATAGATATCATATAAAGTATTTGTAACACTTTGTAAGAA
CCTTGCTACTATAGGTAAAAAGAAACATTTCAAAGTACATGCCCTAATTA
GAAAAAAAGTTATAAAAAAATAATGACTAGGGGCGTGTTTTTTTACTAG
TTTGTATCAAATTATATCAAAATTTAAGGTGGAAAAGAATGACGACCACA
TTAACCAGAAATGTAATTATTTTTTTATTTGGTAATTTTTAATATTTGTT
GTGATCTATGTATTTAAAAGTAAATATCAAACAAGAACATAATCCAAACC
CTAAATTGCAAGTCTCGCCCAATTTCTCTATCACTAGTCCTCACTTACGA
TGGCGTTACGTCGCTCTCTCACTTCCTACAACCCATTGTTGCTACTAATT
ACACTAACGAAAAGTTGAATATCCATATATTTATTTGGATGTGAAATTGA
ACGAATCTCGTCAAATTTTTTATTTTGTTGATGGATTTGAGTGGAAGTTT
AGGCAGAACGGGAATGATGGTCTGCAAGTGGTTATAAACATGGGTGAAGA
TAAAATGGAGTTGTCGCCGTTGTATTATAGATCTCTTAGGGGTTTGATTC
TGAGTTATTACTGTATACGTAGCCTCTTTACAACGACCATTCTTCCAAGT
ACCATTTGATCTTTTTAGAATCCAGTTGTCTGAAACACCCTGATTTGGAT
CAAATATCACCAACAACTCTTAAGAACTGGACTAATTAATTGTTTTCTTG
ATCTTGATAACAAGAGGAAACACGTCACCATATCTTTTATTTTAAATTTG
CTTTTGGTGTATTTTCTTTCTTCCCATTTCTTTCTTGATCTGTTCCAGAT
GGTATTTGGTGTGGATAATTTACACCTGGAGATTGTGAACGATGGGAAGG
GGTATGTGATTTACAGAGGATGTGGCTTGTGGTTGAGGATGGTTTATGGC
TGGCCGAGTCTAATTTATATTTATATAAACAAATAAATATATAAAACAAG
GGTAAAATATGTATTTAAGCGTCCTCTTTTAATGGTGACAATTTTTACAG
TTTACTCTCTTTGTTTTTTAATTGTGATGCCCACGATCGAACTCATTCAT

CCCCCCCCCCTTTTTTTTTTAAAATAAAAAATTAAGAAGGGGTACCACCAT
ATACCCGTGTCAGCTTCTTATTCCCAAGCAGTCAAATAGGGACTTAGGTT
GTATGGAAACAGTTCCGTGACTTGGATGGCAGATAAATTTAGTAAACTTA
ACCCTTCAATTAACCTACCTTTTTCTTATTAACTCAATTTCAAGCTAAAT
TCTGATTCTTGTTTGAAAATAAGTTGCATCTTTATTTTTGCATATTATCT
TGTTGCATAGGATCCTTAGCATCTTTTAATAGTTTATTTGAAGCTGAAAG
ATCCAACTAGTTTTGATCTGTTGGCATTTTCCATCATTTGCAACTGTTTC
TTGAAAAAAAATACCTAAAATCAAAATAACCATTTTCAAATCCAAAATTA
TAAGAGAGAATTGTTAATGGACGTGGAATCATAAATCATTAACACAGTTC
AGTACACAAGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTAT
CAGAGAAAGAGACATTACAAGAAGTCACTGATACTAATATTTCTAATGAT
GTTGTATTATTCCCATCCTGTCTCATGCACTCTTTTCATAACCTCCATAA
ACTTAAATTGGAAAATTATGAAGGAGTGGAGGTGGTGTTTGAGATAGAGA
GTGAGAGTCCAACATGTAGAGAATTGGTAACAACTCACAATAACCAACAA
CAGCCTATTATACTTCCCAACCTCCAGGAATTGTATCTAAGGAATATGGA
CAACACGAGTCATGTGTGGAAGTGCAGCAACTGGAATAAATTCTTCACTC
TTCCAAAACAACAATCAGAATCACCATTCCACAACCTCACAACCATAGAA
ATGAGATGGTGTCATGGCTTTAGGTACTTGTTTTCGCCTCTCATGGCAGA
ACTTCTTTCCAACCTAAAGAAAGTCAAGATACTTGGGTGTGATGGTATTG
AAGAAGTTGTTTCAAACAGAGATGATGAGGATGAAGAAATGACTACATTT
ACATCTACCCACACAACCACCAACTTGTTCCCTCATCTTGATTCTCTCAC
TCTAAAATACATGCACTGTCTGAAGTGTATTGGTGGAGGTGGTGCCAAGG
ATGAGGGGAGCAATGAAATATCTTTCAATAATACCACTACAACTACCGAT
CAATTTAAGGTATGTTTGTACATATTTAATTATATATTTAATTTCCTTGT
TAATTTCCTTTTCTTTGCAATATTCTATGCGAACTCAAGAATGGGATTTG
GAGGCATATAAAGTTACATTCATTTGAACAAGTATTACCTTTTATTTGTT
ATTTATCATTTTCATATCAAGTACCTATAACATTTCTTTTTTATTTTTCT
AATTAGAAGAGGTCCACATGTCTAATTAGGTTTTCCATTCTATGTGTAAC
CTCTATTCTCTCTGTAATCAAGCATCTTAGATTATTTATCCATTTTCATA
ATTGTGTTTATTTTTACAGTTTTTTTTTTTATTTAATTTTAATAATTTAA
TTTTAATTTATTTATTATTTTTTTTTTTGGTAATTGCAACCTGTCATATAT
TCAAGTCTTAATGTAACATAATAATACATTTTATACCCACTATACTAAGA
TAATAATTACCTAAAGGGATGGATGCCATGACACTGCTACACTTCAGNAA
CTCTAGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTTT
TGATGGGTAATATAGGCAATTTAAGTTTTATTTCTGTTAAAGCAGTATTT
AGCTAGTAGTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTCAAAATCT
GGTCATTGTACCCAGAATTTAGTTAAATGTAACATTTTAGATATTAGGGG
TCATCAGGTGACAGATATTGTAGAATAGAACAATATGTAATATTACCCAA
AACTATTTTTTCTAAGGTTGCTCTGTTAAATATGTGCTTTCTTGATTTCA
TTGAATTTGCATTCGTATATTTTAGGTGGTAAACTGATTGTCTCTTCAAT
AAATCCTGAAATTAATTAAAAAAAAAAAAACAAAAGTACATTTTTGATTT
GGAGAGCACTGGTATCATTTAGTATAGAAAAAAACTAGATTTTGAATTAY
CTTTCTTATATAAAAGTTGTGTATATAGTTTAATTAGTTTTACATCATTT

TTCTATGTGTTGTTGCAGTTGTCTGAAGCAGGTGGTGTTTGTTGGAGCTT
ATGCCAATACTCTAGAGAGATAGAGATATATAGGTGTGATGCACTGTCAA
GTGTAATTCCATGTTACGCAGCAGGACAAATGCAAAAGCTGCAAGTGCTG
ACAGTCAGTTCTTGTAATGGTCTGAAGGAGGTATTTGAAACTCAATTAGG
GACGAGCAGCAACAAAAACAACGAGAAGAGTGGTTGTGAGGAAGGAATTC
CAAGAGTAAATAACAATGTTATTATGCTTCCCAATCTAAAGATATTGGAA
ATCTACGGTTGTGGGGGTTTGGAACATATATTCACATTCTCTGCACTTGA
AAGCCTGAGACAGCTCCAAGAGTTAACGATTAAGGGTTACTACTCTTGTC
AATCTTCCAAACCTCAAAGAAATGAGGTTGGAGTGGCTAAGTAATCTGAG
GTATATATGGAAGAGCAATCAGTGGACAGCATTTGAGTTTCCAAACCTAA
CAAGAGTTGAAATTTGTGAATGTAATTCATTAGAACATGTATTTACTAGT
TCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTACATATATTTAACTG
CAGTCTGATGGAGGAGGTAATTGTTAAGGATGCAGATGTTTCTGTAGAAG
AAGACAAAGAGAAAGAATCTGATGGCAAGACGAATAAGGAGATACTTGTG
TTACCTCATCTAAAGTCCTTGAAATTACAACTTCTTCGAAGTCTTAAGGG
GTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATTATTGGATACTTTAG
AAATCAAAAGATGCCCAACAATAACCACCTTCACCAAAGGAAATTCCGCT
ACTCCACAACTAAAAGAAATACAAACAAATTTTGGCTTCTTTTATGCTGC
AGGGGAAAAAGACATCAACTCTCTTATAAAGATCAAACAACAGGTAAATC
AGATCTTTGTTGCTTTAATAATTCTTAAACTACATTTGAAAAGCTTCATG
CAAGTTTTTTTGTTATATTGTCAAAAACCGCAACCTACATTCAGCTTTAT
ATTTATGTACTTTATGCAGGATTTCAAACAAGACTCAGATTAATGTGAAG
TGAATATTAAAGGTAAATTATATTTTCATGTTCCTAGTTGCCTATTAATT
AATGGCCTTTTAGTTCATGATTTTTGGATGTATTCTTCATGATGATGTGA
ATCTTCTAATACCCCATTCATTGTTTGGTTGAATGTTGACTCTATGTCAG
GATGAATATTCAAGGGAAGAATTGTTCATCAWATGAAGGACATTAAAGAA
CATGGATGCTATGAAGATGTTGGGAAAACATATGTATCAAGTGGCAARCT
GCTTAATGATCTAAGTTTGTTGGTTGANGATGTTGATTTTAATATTTCAA
ATTCATTGGTTATATGGGCTTATCAATAGTGTTAATGGGATAATGAGTGA
CTTAACCTAAATTATGTTGTTGGTAAATGTTGGACAAGTATGGAAAATTA
GGAATGACTTGTGAAAAAAAAATAAAAAAAAA **(SEQ ID NO:94)**

RG2D deduced polypeptide sequence (SEQ ID NO:95)

MAMETANEIIKQVVPVLMVPINDYLRYVVSCRKYISDMDLKMKELKEAKDNVEE
HKNHNISNRLEVPAAQVQSWLEDVEKINAKVETVPKDVGCCFNLKIRYRAGRDAF
NIIEEIDSVMRRHSLITWTDHPIPLGRVDSVMASTSTLSTEHNDFQSREVRFSEALKA
LEANHMIALCGMGRVGKTHMMQRLKKVAKEKRKFGYIIEAVIGEISDPIAIQQVVA
DYLCIELKESDKKTRAEKLRQGFKAKSDGGNTKFLIILDDVWQSVDLEDIGLSPSPN
QGVDFKVLLTSRDEHVCSVMGVEANSIINVGLLIEAEAQRLFQQFVETSEPELHKIG
EDIVRRCCGLPIAIKTMACTLRNKRKDAWKDALSRLQHHDIGNVATAVFRTSYENL
PDKETKSVFLMCGLFPEDFNIPTEELMRYGWGLKLFDRVYTIIEARNRLNTCIERLV
QANLLIGSDNGVHVKMHDLVRAFVLGMYSEVEQASIVNHGNMPGWPDENDMIVH

SCKRISLTCKGMIEIPVDLKFPKLTILKLMHGDKSLKFPQEFYEGMEKLQVISYDKM .
KYPLLPLAPQCSTNIRVLHLTECSLKMFDCSSIGNLSNLEVLSFANSRIEWLPSTVRN
LKKLRLLDLRFCDGLRIEQGVLKSLVKLEEFYIGNAYGFIDDNCKDMAERSYNLSA
LEFAFFNNKAEVKNMSFENLERFKISVGCSFDGNISMSSHSYENMLQLVTNKGDVL
DSKLNGLFLKTEVLFLSVHGMNDLEDVEVKSTHPTQSSSFCNLKVRIISKCVELRYL
FKLHVANTLSSLEHLEVCGCENMEELIHTGIGGCGEETITFPKLKSLSLSQLPKLSGL
CHNVNIIGLPHLVDLKLKGIPGFTVIYPQNKLRTSSLLKEEVVIPKLETLQIDGMENL
EEIWPCELSGGEKVKLREIKVSSCDKLVNLFPHNPMSLLHHLEELKVKNCRSIESLF
NIDLDCVSAIGEEDNKSILRRIKVKNLGKLREVWRIKGADNSRPLIHGFPAVESISIW
GCKRFRNIFTPITANFDLVALLEIHIGNYRENHESEEQIEILSEKETLQEVTDTNISND
VVLFPSCLMHSFHNLHKLKLENYEGVEVVFEIESESPTCRELVTTHNNQQQPIILPN
LQELYLRNMDNTSHVWKCSNWNKFFTLPKQQSESPFHNLTTIEMRWCHGFRYLFS
PLMAELLSNLKKVKILGCDGIEEVVSNRDDEDEEMTTFTSTHTTTNLFPHLDSLTLK
YMHCLKCIGGGGAKDEGSNEISFNNTTTTTDQFKLSEAGGVCWSLCQYSREIEIYRC
DALSSVIPCYAAGQMQKLQVLTVSSCNGLKEVFETQLGTSSNKNNEKSGCEEGIPR
VNNNVIMLPNLKILEIYGCGGLEHIFTFSALESLRQLQELTIKGYYTLVNLPNLKEM
RLEWLSNLRYIWKSNQWTAFEFPNLTRVEICECNSLEHVFTSSMVGSLLQLQELHIF
NCSLMEEVIVKDADVSVEEDKEKESDGKTNKEILVLPHLKSLKLQLLRSLKGFSLGK
EDFSFPLLDTLEIKRCPTITTFTKGNSATPQLKEIQTNFGFFYAAGEKDINSLIKIKQQ
DFKQDSD.CEVNIK

RG2E polynucleotide sequence (SEQ ID NO:96)

TGGGAAGACACAATGATGCAAAGGTTGAAGAAGGTTGCTAAAGAAAATAGAAT
GTTCAATTATATGGTTGAGGCAGTTATAGGGGAAAAGACAGACCCACTTGCTAT
TCAACAAGCTGTAGCGGATTACCTTTGTATAGAGTTAAAAGAAAGCACTAAACC
AGCAAGAGCTGATAAGCTTCGTGAATGGTTTAAGGCCAACTCTGGAGAAGGTA
AGAATAAGTTCCTTGTAATATTTGATGATGTTTGGCAGTCCGTTGATCTGGAAG
ACATTGGTTTAAGTCATTTTCCAAATCAAGGTGTCGACTTCAAGGTCTTGTTGA
CTTCACGAGACGAACATGTTTGCACAGTAATGGGGGTTGAAGCTAATTCAATTC
TTAATGTGGGACTTCTAGTAGAAGCAGAAGCACAAAGTTTGTTCCAGCAATTTG
TAGAAACTTTTGAGCCCGAGCTCCATAAGATAGGAGAAGATATCGTAAGGAAG
TGTTGTGGTTTACCTATTGCCATTAAAACCATGGCATGTACTCTAAGAAATAAA
AGAAAGGATGCATGGAAGGATGCACTTTTGCATTTAGAGTACCATGACATTAGC
AGTGTTGCGCCCAAAGTCTTTGAAACGAGCTACCATAATCTCCACAACAAGGAG
ACTAAATCTGTGTTTTTGATGTGTGGTTTTTTTCCTGAAGACTTCAATATTCCAA
TCGAGGAGTTGATGAGGTATGGATGGGGCTTAAAGATATTTGATAGAGTTTATA
CTATTAGACAAGCAAGAATCAGGCTCAACACCTGCATTGAGCGACTGGTGCAG
ACAAATTTGTTAATAGAAAGTGATGATGGTGTGCACGTCAAGATGCATGATCTG
GTCCGTGCTTTCGTTTTGGTTATGTTTTCTGAAGTTGAACATGCTTCAATTATCA
ACCATGGTAATATGCTTGGATGGCCTGAAAATTATATGACCAACTCTTGCAAAA
CAATTTCATTAACATGCAAGAGTATGTCTGAATTTCCGGGAGATCTCAAGTTTC
CAAACCTAACGATTTTGAAACTCATGCATGGAGATAAGTTGCTAAGATATCCTC

AAGACTTTTATGAAGGAATGGAAAAGCTCTGGGTTATATCATATGATGAAATGA
AGTATCCATTGCTTCCCTCGTTACCTCAATGCTCCATCAACCTTCGAGTGCTTCA
CCTCCATCGATGCTCATTAATGATGTTTGATTGCTCTTGTATTGGAAATATGTTG
AATCTGGAAGTGCTTAGCTTTGTTAAATCTGGCATTGAATGGTTACCTTCCACA
ATAGGAAATTTAAAGAAGCTAAGGTTACTTGATCTGAGAGATTGTTATGGTCTT
CGTATAGAAAAAGGTGTCTTGAAAAATTTGGTGAAAATTGGAGGAATTTATATT
GGTAGAGCAGATATTTTATAGAT

RG2E deduced polypeptide sequence (SEQ ID NO:97)

WEDTMMQRLKKVAKENRMFNYMVEAVIGEKTDPLAIQQAVADYLCIELKESTKP
ARADKLREWFKANSGEGKNKFLVIFDDVWQSVDLEDIGLSHFPNQGVDFKVLLTS
RDEHVCTVMGVEANSILNVGLLVEAEAQSLFQQFVETFEPELHKIGEDIVRKCCGL
PIAIKTMACTLRNKRKDAWKDALLHLEYHDISSVAPKVFETSYHNLHNKETKSVFL
MCGFFPEDFNIPIEELMRYGWGLKIFDRVYTIRQARIRLNTCIERLVQTNLLIESDDG
VHVKMHDLVRAFVLVMFSEVEHASIINHGNMLGWPENYMTNSCKTISLTCKSMSE
FPGDLKFPNLTILKLMHGDKLLRYPQDFYEGMEKLWVISYDEMKYPLLPSLPQCSI
NLRVLHLHRCSLMMFDCSCIGNMLNLEVLSFVKSGIEWLPSTIGNLKKLRLLDLRD
CYGLRIEKGVLKNLVKIGGIYIGRADIL.

RG2F polynucleotide sequence (SEQ ID NO:98)

CTGTGGAAGACACAATGATGCAAAGGCTGAAAAAGGTTGTGCATGAAAAGAAA
ATGTTTAACTTTATTGTTGAAGCAGTTATAGGGGAAAAGACAGACCCCGTTGCC
ATTCAGGATGCTATAGCAGATTACCTAGGTGTAGAGCTCAATGAAAAATCTAAG
CAAGCAAGAGCTGATAAGCTCCGTCAAGGATTCAAGGACAAATCAGATGGAGG
CAAAAATAAGTTCTTTGTAATACTTGACGATGTTTGGCAGTCTGTTGATCTGGA
AGATATTGGTTTAAGTCCTTTTCCAAATCAAGGCGTCGACTTCAAGGTCTTGTT
GACATCACGAGACAGACATGTTTGCACAGTGATGGGGGTTGAAGCCAAATTAA
TTCTAAACGTGGGACTTCTAATTGAAGCTGAAGCACAAAGTTTGTTCCACCAAT
TTGTTGTCACTTCTGAGCCCGAGCTCCATAAGATAGGAGAAGATATTGTAAAGA
AGTGTTTCGGTCTGCCAATTGCCATCAAAACCATGGCATGTACTCTACGACATA
AAAGAAAGGATGCATGGAAGGATGCACTTTCACGTTTAGAGCACCATGACATT
CAAAGTGTTGTGCCTAAAGTATTTGAAACGAGCTACAACAATCTCAAAGACAA
GGAGACTAAATCCGTATTTTTGATGTGTGGTTTGTTTCCTGAAGACTTGGATAT
ACCTATCGAGGAGTTGATGAGGTATGGATGGGGCTTAAGATTATTTGATAGAGT
TAATACTATTACACAAGCAAGAAACAGGCTCAACACCTGCATTGAGCGACTGG
TGCACACAAATTTGTTAATTGAAAGTGTTGATGGTGTGCATGTCAAGATGCATG
ATCTGGTTCGTGCTTTTGTTTTGGGAATGTTTTCTGAAGTGGAGCATGCTTCAAT
TGTCAACCATGGTAATATGCCCGAGTGGACTGAAAATGATATGACTGACTCTTG
CAAACAAATTTCATTAACATGCAAGAGTATGTTGGAGTTTCCTGGAGACCTCAA
GTTTCCAAACCTAAAGATTTTGAAACTTATGCATGGAGGTAAGTCACTAAGGTA
TCCTCAAGACTTTTATCAAGGAATGGAAAAGCTGGAGGTTATATCATACGATGA
AATGAAGTATCCATTGCTTCCCTCGTTGCCTCAATGTTCCACCATCCTTCGAGTG

CTTCATCTCCATGAATGTTCATTAAGGATGTTTGATTGCTCTTCAATCGGTAATC.
TTTTCAACATGGAAGTGCTCAGCTTTGCTAATTCTAGCATTGAATTGTTACCTTC
CGTAATTGGAAATTTGAAGAAGTTGCGGCTGCTAGATTTGACAAACTGTTATGG
TGTTCGTATAGAAAAGGATGTCTTGAAAAATTTGGTGAAACTTGAAGAGCTTTA
TATTAGGAATGGTCTACCAGTTTACAGAGGAT

RG2F deduced polypeptide sequence (SEQ ID NO:99)

VEDTMMQRLKKVVHEKKMFNFIVEAVIGEKTDPVAIQDAIADYLGVELNEKSKQA
RADKLRQGFKDKSDGGKNKFFVILDDVWQSVDLEDIGLSPFPNQGVDFKVLLTSRD
RHVCTVMGVEAKLILNVGLLIEAEAQSLFHQFVVTSEPELHKIGEDIVKKCFGLPIAI
KTMACTLRHKRKDAWKDALSRLEHHDIQSVVPKVFETSYNNLKDKETKSVFLMCG
LFPEDLDIPIEELMRYGWGLRLFDRVNTITQARNRLNTCIERLVHTNLLIESVDGVH
VKMHDLVRAFVLGMFSEVEHASIVNHGNMPEWTENDMTDSCKQISLTCKSMLEFP
GDLKFPNLKILKLMHGGKSLRYPQDFYQGMEKLEVISYDEMKYPLLPSLPQCSTILR
VLHLHECSLRMFDCSSIGNLFNMEVLSFANSSIELLPSVIGNLKKLRLLDLTNCYGV
RIEKDVLKNLVKLEELYIRNGLPVYRG

RG2G polynucleotide sequence (SEQ ID NO:100)

GAAGACACGATGATGAAGAACTGAAGGAGGTCGTGGGACAAAAGAAATCATTC
AATATTATTATTCAAGTGGTCATAGGAGAGAAGACAAACCCTATTGCAATTCAG
CAAGCTGTAGCAGATTACCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGC
AAGAGCTGATAAGCTTCGTAAACGGTTTGAAGCCGATGGAGGAAAGAATAAGT
TCCTTGTAATACTTGACGATGTATGGCAGTTTGTCGATCTTGAAGATATTGGTTT
AAGTCCTCTGCCAAATAAAGGTGTCAACTTCAAGGTCTTGTTGACGTCAAGAGA
TTCACATGTTTGCACTCTGATGGGAGCTGAAGCAAATTCAATTCTTAATATAAA
AGTTTTAAAAGATGTAGAAGGACAAAGTTTGTTCCGCCAGTTTGCTAAAAATGC
GGGTGATGATGACCTGGATCCTGCTTTCAATGGGATAGCAGATAGTATTGCAAG
TAGATGTCAAGGTTTGCCCATTGCCATCAAAACCATTGCCTTAAGTCTTAAAGG
TAGAAGCAAGTCTGCATGGGACGTTGCACTTTCTCGTCTGGAGAATCATAAGAT
TGGTAGTGAAGAAGTTGTGCGTGAAGTTTTTAAAATTAGCTACGACAATCTCCA
AGATGAGGTTACTAAATCTATTTTTTTACTTTGTGCTTTATTTCCTGAAGATTTT
GATATTCCTACTGAGGAGTTGGTGAGGTATGGGTGGGGCTTGAAATTATTTATA
GAAGCAAAAACTATAAGAGAAGCAAGAAACAGGCTCAACACCTGCACTGAGCG
GCTTAGGGAGACAAATTTGTTATTTGGAAGTGATGACATTGGATGTGTCAAGAT
GCACGATGTGGTGCGTGATTTTGTTTTGCATATATTCTCAGAAGTCCAACACGC
TTCAATTGTCAACCATGGTAACGTGTCAGAGTGGCTAGAGGAAAATCATAGCAT
CTACTCTTGTAAAAGAATTTCATTAACATGCAAGGGTATGTCTCAGTTTCCCAA
AGACCTCAAATTTCCAAACCTTTCAATTTTGAAACTTATGCATGGAGATAAGTC
ACTGAGCTTTCCTGAAAACTTTTATGGAAAGATGGAAAAGGTTCAGGTAATATC
ATATGATAAATTGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCCACCAA
CGTTCGAGTGCTTCATCTTCATTACTGTTCATTAAGGATGTTTGATTGCTCTTCA
ATTGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTAATTCTAACATTGAA

TGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGCTACTAGATTTGACA
AATTGTAAAGGTCTTCGTATAGATAATGGTGTCTTAAAAAATTTGGTCAAACTT
GAAGAGCTTTATATGGGTGTTAATCGTCCGTATGGACAGGCCGTTAGCTTGACA
GATGAAAA

RG2G deduced polypeptide sequence (SEQ ID NO:101)

RHDDEELKEVVGQKKSFNIIIQVVIGEKTNPIAIQQAVADYLSIELKENTKEARADKL
RKRFEADGGKNKFLVILDDVWQFVDLEDIGLSPLPNKGVNFKVLLTSRDSHVCTL
MGAEANSILNIKVLKDVEGQSLFRQFAKNAGDDDLDPAFNGIADSIASRCQGLPIAI
KTIALSLKGRSKSAWDVALSRLENHKIGSEEVVREVFKISYDNLQDEVTKSIFLLCAL
FPEDFDIPTEELVRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVK
MHDVVRDFVLHIFSEVQHASIVNHGNVSEWLEENHSIYSCKRISLTCKGMSQFPKDL
KFPNLSILKLMHGDKSLSFPENFYGKMEKVQVISYDKLMYPLLPSSLECSTNVRVLH
LHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKLRLLDLTNCKGLRID
NGVLKNLVKLEELYMGVNRPYGQAVSLTDE

RG2H polynucleotide sequence (SEQ ID NO:102)

TGAAGGAGGTTGTGGAACGAAAGAAAATGTTCAGTATTATTGTTCAAGTG
GTCATAGGAGAGAAGACAAACCCTATTGCTATTCAGCAAGCTGTAGCAGA
TTACCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGCAAGAGCTGATA
AGCTTCGTAAATGGTTCGAGGCCGATGGAGGAAAGAATAAGTTCCTTGTA
ATACTTGACGATGTATGGCAGTTTGTCGATCTTGAAGATATTGGTTTAAG
TCCTCTGCCAAATAAAGGTGTCAACTTCAAGGTCTTGTTGACGTCAAGAG
ATTCACATGTTTGCACTCTGATGGGAGCCGAAGCCAATTCAATTCTCAAT
ATAAAAGTTTTAACAGCTGTAGAAGGACAAAGTTTGTTCCGCCAGTTTGC
TAAAAATGCGGGTGATGATGACCTGGATCCTGCTTTCAATAGGATAGCAG
ATAGTATTGCAAGTAGATGTCAAGGTTTGCCCATTGCCATCAAAACCATT
GCCTTAAGTCTTAAAGGTAGAAGCAAGCCTGCGTGGGACCATGCGCTTTC
TCGTTTGGAGAACCATAAGATTGGTAGTGAAGAAGTTGTGCGTGAAGTTT
TTAAAATTAGCTATGACAATCTCCAAGATGAGATTACTAAATCTATTTTT
TTACTTTGTGCTTTATTTCCTGAAGATTTTGATATTCCTACTGAGGAGTT
GATGAGGTATGGATGGGGCTTGAAATTATTTATAGAAGCAAAAACTATAA
GAGAAGCAAGAAACAGGCTCAACACCTGCACTGAGCGGCTTAGGGAGACA
AATTTGTTATTTGGAAGCGATGACATTGGATGCGTCAAGATGCACGATGT
GGTGCGTGATTTTGTTTTGCATATATTCTCAGAAGTCCAGCACGCTTCAA
TTGTCAACCATGGTAACGTGTCAGAGTGGCTAGAGGAAAATCATAGCATC
TACTCTTGTAAAAGAATTTCATTAACATGCAAGGGTATGTCTGAGTTTCC
CAAAGACCTCAAATTTCCAAACCTTTCAATTTTGAAACTTATGCATGGAG
ATAAGTCGCTGAGCTTTCCTGAAAACTTTTATGGAAAGATGGAAAAGGTT
CAGGTAATATCATATGATAAATTGATGTATCCATTGCTTCCCTCATCACT
TGAATGCTCCACTAACGTTCGAGTGCTTCATCTCCATTATTGTTCATTAA
GGATGTTTGATTGCTCTTCAATTGGTAATCTTCTCAACATGGAAGTGCTC

AGCTTTGCTAATTCTAACATTGAATGGTTACCATCTACAATTGGAAATTT
GAAGAAGCTAAGGCTACTAGATTTGACAAATTGTAAAGGTCTTCGTATAG
ATAATGGTGTCTTAAAAAATTTGGTCAAACTTGAAGAGCTTTATATGGGT
GTTAATCATCCGTATGGAC

RG2H deduced polypeptide sequence (SEQ ID NO:103)

KEVVERKKMFSIIVQVVIGEKTNPIAIQQAVADYLSIELKENTKEARADKLRKWFEA
DGGKNKFLVILDDVWQFVDLEDIGLSPLPNKGVNFKVLLTSRDSHVCTLMGAEAN
SILNIKVLTAVEGQSLFRQFAKNAGDDDLDPAFNRIADSIASRCQGLPIAIKTIALSLK
GRSKPAWDHALSRLENHKIGSEEVVREVFKISYDNLQDEITKSIFLLCALFPEDFDIP
TEELMRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVKMHDVVR
DFVLHIFSEVQHASIVNHGNVSEWLEENHSIYSCKRISLTCKGMSEFPKDLKFPNLSI
LKLMHGDKSLSFPENFYGKMEKVQVISYDKLMYPLLPSSLECSTNVRVLHLHYCSL
RMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKLRLLDLTNCKGLRIDNGVLKN
LVKLEELYMGVNHPYG

RG2I polynucleotide sequence (SEQ ID NO:104)

AAGAAGAGCTGAAGGAGGTTGTGGAACAAAAGAAAACGTTCAATATTATT
GTTCAAGTGGTCATAGGAGAGAAGACAAACCCTATTGCTATTCAGCAAGC
TGTAGCAGATTCCCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGCAA
GAGCTGATAAGCTTCGTAAATGGTTCGAGGCTGATGGAGGAAAGAATAAG
TTCCTCGTNATACTTGACGATGTATGGCNGTTTGTTGATCTTGAAGATAT
TGGTTTAAGTCCTCATCCAAATAAAGGTGTCANCTTCAAGGTCTTGTTGA
CGTCAAGAGATTCACATGTTTGCACTCTGATGGGAGCTGAAGCCAATTCA
ATTCTCAATATAAAAGTTTTAAAAGATGTAGAAGGAAAAAGTTTGTTCCG
CCAGTTTGCTAAAAATGCGGGTGATGATGACCTGGATCCTGCTTTCATTG
GGATAGCAGATAGTATTGCAAGTAGATGTCAAGGTTTGCCCATTGCCATC
AAAACCATTGCCTTAAGTCTTAAAGGTAGAAGCAAGTCTGCATGGGACGT
TGCACTTTCTCGTCTGGAGAATCATAAGATTGGTAGTGAAGAAGTTGTGC
GTGAAGTTTTTAAAATTAGCTATGACAATCTCCAAGATGAGGTTACTAAA
TCTATTTTTTTACTTTGTGCTTTATTTCCTGAAGATTTTGATATTCCTAC
TGAGGAGTTGGTGAGGTATGGGTGGGGCTTGAAATTATTTATAGAAGCAA
AAACTATAAGAGAAGCAAGAAACAGGCTCAACACCTGCACTGAGCGGCTT
AGGGAGACAAATTTGTTATTTGGAAGTGATGACATTGGATGCGTCAAGAT
GCACGATGTGGTGCGTGATTTTGTTTTGCATATATTCTCAGAAGTCCAGC
ACGCTTCAATTGTCAACCATGGTAATGTGTCAGAGTGGCTAGAGGAAAAT
CATAGCATCTACTCTTGTAAAAGAATTTCATTAACATGCAAGGGTATGTC
TGAGTTTCCCAAAGACCTCAAATTTCCAAACCTTTCAATTTTGAAACTTA
TGCATGGAGATAAGTCGCTGAGCTTTCCTGAAAACTTTTATGGAAAGATG
GAAAAGGTTCAGGTAATATCATATGATAAATTGATGTATCCATTGCTTCC
CTCATCACTTGAATGCTCCACCAACCTTCGAGTGCTTCATCTCCATGAAT
GTTCATTAAGGATGTTTGATTGCTCTTCAATTGGTAATCTTCTCAACATG

GAAGTGCTCAGCTTTGCTAATTCTGGCATTGAATGGTTACCATCTACAAT
TGGAAATTTGAAGAAGCTAAGGCTACTGGATCTGACAGATTGTGGAGGTC
TTCATATAGATAATGGCGTCTTAAAAAATTTGGTCAAACTTGAAGAGCTT
TATATGGGTGCTAATCGTCTGTTTGGAAAGTGCCAT

RG2I deduced polypeptide sequence (SEQ ID NO:105)

EELKEVVEQKKTFNIIVQVVIGEKTNPIAIQQAVADSLSIELKENTKEARADKLRKWF
EADGGKNKFLVILDDVW?FVDLEDIGLSPHPNKGV?FKVLLTSRDSHVCTLMGAEA
NSILNIKVLKDVEGKSLFRQFAKNAGDDDLDPAFIGIADSIASRCQGLPIAIKTIALSL
KGRSKSAWDVALSRLENHKIGSEEVVREVFKISYDNLQDEVTKSIFLLCALFPEDFDI
PTEELVRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVKMHDVV
RDFVLHIFSEVQHASIVNHGNVSEWLEENHSIYSCKRISLTCKGMSEFPKDLKFPNLS
ILKLMHGDKSLSFPENFYGKMEKVQVISYDKLMYPLLPSSLECSTNLRVLHLHECSL
RMFDCSSIGNLLNMEVLSFANSGIEWLPSTIGNLKKLRLLDLTDCGGLHIDNGVLKN
LVKLEELYMGANRLFGKCH

RG2J polynucleotide sequence (SEQ ID NO:106) and (SEQ ID NO:107)

ATGTCCGACCCAACAGGGATTGTTGGTGCCATTATTAACCCAATTGCTCA
AACGGCCTTGGTTCCCCTTACAGACCATGTAGGCTACATGATTTCCTGCA
GAAAATATGTGAGGGACATGCAAATGAAAATGACAGAGTTAAATACCTCA
AGAATCAGTGCAGAGGAACACATTAGCCGGAACACAAGAAATCATCTTCA
GATTCCATCTCAAATTAAGGATTGGTTGGACCAAGTAGAAGGGATCAGAG
CGAATGTTGCAAACTTTCCAATTGATGTCATCAGTTGTTGTAGTCTCAGG
ATCAGGCACAAGCTTGGACAGAAAGCCTTCAAGATAACTGAGCAGATCGA
AAGTCTAACGAGACAAAATTCGCTGATTATCTGGACTGATGAACCTGTTC
CCCTGGGAAGAGTTGGTTCCATGATTGCATCCACCTCTGCAGCATCAAGT
GATCATCATGATGTCTTCCCTTCAAGAGAGCAAATTTTTAGGAAAGCACT
AGAAGCACTTGAACCCGTCCAAAAATCCCACATAATAGCCTTATGGGGGA
TGGGCGGAGTGGGGAAGACCACGATGATGAAGAAGCTGAAAGAGGTCGTG
GAACAAAAGAAAACGTGCAATATTATTGTTCAAGTGGTCATAGGAGAGAA
GACAAACCCTATTGCTATCCAGCAAGCTGTAGCAGATTACCTCTCTATAG
AGCTGAAAGAAAACACTAAAGAAGCAAGAGCTGATAAGCTTCGTAAACGG
TTCGAAGCCGATGGAGGAAAGAATAAGTTCCTTGTAATACTTGACGATGT
ATGGCAGTTTTTCGATCTTGAAGATATTGGTTTAAGTCCTCTGCCAAATA
AAGGTGTCAACTTCAAGGTCTTGTTGACGTCAAGAGATTCACATGTTTGC
ACTCTGATGGGAGCTGAAGCCAATTCTATTCTCAATATAAAAGTTTTAAA
AGATGTAGAAGGAAAAAGTTTGTTCCGCCAGTTTGCTAAAAATGCGGGTG
ATGATGACCTGGATCCTGCTTTCATTGGGATAGCAGATAGTATTGCAAGT
AGATGTCAAGGTTTGCCCATTGCCATCAAAACCATTGCCTTAAGTCTTAA
AGGTAGAAGCAAGTCTGCATGGGACGTCGCACTTTCTCGTCTGGAGAATC
ATAAGATTGGTAGTGAAGAAGTTGTGCGTGAAGTTTTAAAATTAGCTAT
GACAATCTCCAAGATGAGGTTACTAAATCTATTTTTTTACTCTGTGCTTT

ATTTCCTGAAGATTTTGATATTCCTATTGAGGAGTTGGTGAGGTATGGGT
GGGGCTTGAAATTATTTATAGAAGCAAAAACTATAAGAGAAGCAAGAAAC
AGGCTCAACAACTGCACTGAGCGGCTTAGGGAGACAAATTTGTTATTTGG
AAGTCATGACTTTGGGTGCGTCAAGATGCACGATGTGGTGCGTGATTTTG
TTTTGCATATGTTTTCAGAAGTCAAGCATGCTTCAATTGTCAACCATGGT
AACATGTCAGAGTGGCCAGAGAAAAATGATACCAGCAACTCTTGTAAAAG
AATTTCATTAACATGCAAGGGTATGTCTAAGTTTCCTAAAGACATCAACT
ATCCAAACCTTTTGATTTTGAAACTTATGCATGGAGATAAGTCGCTGTGC
TTTCCTGAAAACTTTTATGGAAAGATGGAAAAGGTTCAGGTAATATCATA
TGATAAATTGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCCACTA
ACGTTCGAGTGCTTCATCTCCATTATTGTTCATTAAGGATGTTTGATTGC
TCTTCAATTGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTAATTC
TAACATTGAATGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGC
TACTAGATTTGACAAATTGTAAAGGTCTTCGTATAGATAATGGTGTCTTA
AAAAATTTGGTCAAACTTGAAGAGCTTTATATGGGTGTTAATCGTCCGTA
TGGACAGGCCGTTAGCTTGACAGATGAAAACTGCAATGAAATGGTAGAAG
GTTCCAAAAAACTTCTTGCACTAGAATATGAGTTGTTTAAATACAATGCT
CAAGTGAAGAATATATCCTTCGAGAATCTTAAACGATTCAAGATCTCAGT
GGGATGTTCTTTACATGGATCTTTCAGTAAAAGCAGGCACTCATACGAAA
ACACGTTGAAGTTGGCCATTGACAAAGGCGAACTATTGGAATCCCGAATG
AACGGGTTGTTTGAGAAAACGGAGGTTCTTTGTTTAAGTGTGGGGGATAT
GTATCATCTTTCAGATGTTAAGGTGAAGTCCTCTTCGTTCTACAATTTAA
GAGTCCTTGTCGTTTCAGAGTGTGCAGAGTTGAAACACCTCTTCACACTT
GGTGTTGCAAATACTTTGTCAAAGCTTGAGCATCTTAAAGTCTACAAATG
CGATAATATGGAAGAACTCATACATACCGGGGGTAGTGAAGGAGATACAA
TTACATTCCCCAAGCTGAAGCTTTTATATTTGCATGGGCTGCCAAACCTA
TTGGGTTTGTGTCTTAATGTCAACGCAATTGAGCTACCAAAACTTGTGCA
AATGAAGCTTTACAGCATTCCGGGTTTCACAAGCATTTATCCGCGGAACA
AGTTGGAAGCATCTAGTTTGTTGAAAGAAGAGGTACATATACATATAGTT
TATGTTAATACATTTTAAACAATCTTTTCAACTAAAAGTTTCAGAATATA
TCTGTATTTTGATTGTATGATGTGTTAGTGTTTGGATGTGGCTATTAAAG
GATAATTATTTGGCAGGTTGTGATTCCTAAGTTGGATATACTTGAAATTC
ATGACATGGAGAATTTAAAGGAAATATGGCCTAGTGAGCTTAGTAGAGGT
GAGAAAGTTAAGTTGAGAAAGATTAAAGTGAGAAATTGTGATAAACTTGT
GAATCTATTTCCACACAATCCCATGTCTCTGCTGCATCATCTTGAAGAGC
TTATAGTCGAGAAATGTGGTTCCATTGAAGAGTTGTTCAACATCGACTTG
GATTGTGCCAGTGTAATTGGAGAAGAAGACAACAACAGCAGCTTAAGAAA
CATCAATGTGGAGAATTCAATGAAGCTAAGAGAGGTGTGGAGGATAAAAG
GTGCAGATAACTCTCGTCCCCTCTTTCGTGGCTTTCAAGTTGTTGAAAAG
ATAATCATTACGAGATGTAAGAGGTTTACAAATGTATTCACACCTATCAC
CACAAATTTTGATCTGGGGGCACTTTTGGAGATTTCAGTTGATTGTAGAG
GAAATGATGAATCAGACCAAAGTAACCAAGAGCAAGAGCAGGTATGGATT
TCAATTTTACTCTTTTACTTAATTAATGATTAAGCCCCTGCTTTTTAATA

AAAAGGGGACAAACCATTTCTTGACTTAATGTTGCAATACAAGTCATGTA
TAAGAGTGATTAACTTTTTTTATTTATAAAATAACTACAAAACATGTTT
TTTCATTATAGATCATGTATAAATGTGACTAATTTTTTCATCGCCTAAC
TTTTGTTGATAAATCATTAGAAATGTCACTAATTACTTTTAGTATTTAT
AAAATAACTACAAAACATGTTTTTCATTATAGATCATGTATATATCAAC
TAAAAATATTATTCCCTTACACAAAAAAAAAGGTTCAAGAAAGCCTGTA
TTTCGAAATAACTAAAAAGAAAATATTTGATATTCACTAAGAGAAATTTT
TTTCTAAACATGATCGCAAATGATTAAAACTTAAATTAAAACTAAAAAGA
TTTTTATATATGTTATNCAAAATTAAAATTTGAAATTAAGTTTATAATTC
TNGTNTCACAAAGGGATATATATAGTAAAATATTATTTTTTTGCAGTCAT
GCATAGTTGTATTTTTAAATGATTTATTAACGTGGTAGGAGTGGAAACCA
CTCAATCTAGTAGACCCACTATCACATGTCACATCAGCTTTACATCTATT
TTTCTTTCTCCTTTTTTCATCTTTTTAAACTCATAACACNTAAAANTANC
ATATTTTCCAACACACTNAACTCATTGTCACATTATTATTTTTAATTTAA
TTAAATTNGAAAATTAAAATTAANTAAANCNTAACATTTTTTAATTAAAA
AATATTAATCCAAATAAAAANTNCACGATAAATTAAAAANGTTTANTTTG
GAAAAAAANCC **(SEQ ID NO:106)**

Sequence gap

ATAACCCTTTCAAGGGTCAACTCAAGTCCAAGTTAAAGTCAAGGTCAAAA
CCTTGGTTAAAGTCAACTTTGGTCAAAGTCAACATCTACTTGACTCACCT
CACCGAGTTGGTCCACCAACTTGTCGAGTCCCTTAATCCACAAACTTCAA
GAACTTCGATCCTACTCGTCGAGTCTTTCAAGAACTCTTCGAGTTTCCAT
TACACAGAATCGGGACCTTTTGCTCATGACTCGCCGAGTTCATCCTTGAA
CTTGTCGAGTCTAGCTTCATACGAGTTCGAGTGTTTAGTCCTTGACTCGT
CGAGTTCTTCCTTGAACTCGTCGAGTCCATCTTCGTATAGTTGGGACATT
GCCTTGAACTCACCGAGTTCATCATTGAACTCATCGAGTCCTTCGATCTT
CAAGTCCATAATCCTGTCCATCTTGTTGAGTCCTCTTCTAGACTCAACCA
GATTCCTCAGAAACAGAAAAGGTTAGGGAACCATTACCTGACTCGCCGAG
TCCCAAGAACGAATCCCCGAGTCCCCCAATGTCCATGACCATACAATCGA
TTTTCGTTGGGCTCATTGCATCCAAAGCATAGATCTAACCTCCTAGGGTC
CATATTACACGTAAAGCTACGAACTTGACGTCCATGCATGGGGGATTTGG
CTCAAATGGCATTAAAATGGGGTTTATCTGATGCATGGGACTCCCATGGC
CATAAAGTTAACACCTTTATGCCATGGGAATCCTCAATGGTTCCATATCT
GAAGTTAACACTCTACAATATGTTCTAAACCCGAAGGTGGCTTAGAAATG
CCCCAAAATGGCAAGATTCAAGCCTTAAAGGAGATCTAACAAATGATAAG
TCAAGGTTCAAGCTTTTTACCTTGAATAAGCTGGAAATGAAGCAAAATCT
CTGGATCCACTTGCTTCTTCAAGAACCCCCAAGCTTCCACTTCTTCCTTC
AAGTTTCAAACAACTTTAAACACTCAAAAATGGCTCAAGAACACTCAAAA
AGCTTTAGGGTTTCGAGTTAGGGCTTTTTGGAAGCGAGAGGGACGATGGG
GGCTGAAATGAGGCTAGAAAAAGTGTTTAAATAGGGGGCAAACCCTAAAT
ATTAGGGTTTCATCCAGGCAGCCCTACTCGTCGAGTCGGGCTCCCGACTC
GTCGAGTAGGTCACTTAAAACCCGCGTCCATAATCCAGTCTACTCGACGA
GTTGGGCCTCCAACTCGTCGATTCCGAGTGCAAAACGTTCAATTACTTAA

ATTTAAATATGTACCAGGAACCGGGTGTTACAGTTGAGACTTTATACCTC
CATAAGATAGATCTAGGTGCACATAGCCTGGATCCACAAGCTCCATGTCA
ACAAGCGACTCTTCAAGAAGTTCATTCTTCCTCCTTAAGCACCAAAAAAC
ACACAAAATCACCATGAAGCTCAAGAAATACTCAAATAGAGGATAGGGTT
TCGTTCGTAGGGTTAGAGAGGATGGAGGCTAGAGGAAATGAGGGATAGAG
GCGAGTTAAGGTCTTTAAATAGGGTCCAAGACCCTAAATTAGGGTTTTAA
TCTGGCCAGACGAACGCAGGGTGTTCCCAAATGCATATGTGTCCAAATTC
TCGTGTGCGCCATGCGTACCTCCCTTGTACGCCATGTGTACCGGGTTTGG
TCCAAACCCTTCTAACTTCAAATGATCATAACTTGCACCCCTTATCTGTT
TTCGATGTTCTTTATATCCACGGAAAGGTAACAAGAAGCCCTATACTTCT
ATAAACTTTATTTAATCTGAAAACCAACCGAAATTAAATCCAAAATTCAT
AAAAGTCCCGAACCAACACATTTACCGATACCCTTGGGCTCCAAAACACA
AATTGAAAACCCGGATCATCCAAACTACATCATCCACCTCCAAATGAGCC
CAAACTCAATTATTCAAGGGTTCTAAGCCTGTTAATGCCCACTCCTCGAT
TACCACCCCGCAATGGGAAACGATTCAAAACAGGGCGTTACATAATTTGT
TGTGGTTTTGTATTTTTTATTTCCGGTGAAGGTGAAAGATCCAACTATTT
TTAATCTGTTGGCATTTTCCATCATTTGCAACTGTTTCTTGAAAAAAAAA
TACCTAAAATCAAAATAACCATTTTCAAATCCAAAATTATAAGAGAGAAT
TGTAAATGGACATGGAATCTTAAATCATTAACACAGTTCAGTACACAAGT
TGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTATCAGAGAAAGAG
ACATTACAAGAAGCCACTGACAGTATTTCTAATGTTGTATTCCCATCCTG
TCTCATGCACTCTTTTCATAACCTCCAGAAACTTATATTGAACAGAGTTA
AAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTGAGAGTCCAACAAGTAGA
GAATTGGTAACAACTCACCATAACCAACAACAGCCTGTTATATTTCCCAA
CCTCCAGCATTTGGATCTAAGGGGTATGGACAACATGATTCGCGTGTGGA
AGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAA
TCCCCATTCCACAACCTCACAACCATAAATATTGATTTTGCAGAAGCAT
TAAGTACTTGTTTTCACCTCTCATGGCAGAACTTCTTTCCAACCTAAAGA
AAGTCAATATAAAATGGTGTTATGGTATTGAAGAAGTTGTTTCAAACAGA
GATGATGAGGATGAAGAAATGACTACATTTACATCTACCCACACAACCAC
CATCTTGTTCCCTCATCTTGATTCTCTCACTCTAAGTTTCCTGGAGAATC
TGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGGGGAGCAATGAAATA
TCTTTCAATAATACCACTGCAACTACTGCTGTTCTTGATCAATTTGAGGT
ATGCTTTGTTCATATTCAATTATTTATTTAATTTCCTTTTTTATTTGCAA
TATTCTATAAATAATACATTTTATACCCACTATACTAAGATAATAATTAC
CTAGAGGGATGGATGCTATGACACAGCTGCTACACTTCAGAAACTCTAGT
AAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTTTTGATGGG
TAATATAGGCAATTTAAGTTTTATTTCTGTTAAAGCAGTATTTAGCAAGT
ACTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTGAAAATCTGGTCATT
GTACCCAGAATTTAGTTAAATGTAACATTTTAGATATCAGGGGACATCAG
GTGACAGATATTGTAGAATAGAACAATATATAATATTACCCAAAACTATT
TTTTCTAAGGTTTTTCTGTTAAATATGTGCTTTCTTGATTTCATTGAATT
TGCATTCCTATATTTTAGGTGGTAAAGTGATTGTCTCTTCAATAAATCCC

GAAATTAATTAAAAAAAAAAAAAAAACAAAAGTAAATTTTTGATATGGAGA
GCACTGGTATCATTTAGTATATAAAAAAACTAGATTTTGAATTAAGTTTC
TTATATAAAAGCTGTGTATATAGTTTAATTAGTTTTACATCATTTTTCCA
TGTGGTGTTGCAGTTGTCTGAAGCAGGTGGTGTTTCTTGGAGCTTATGCC
AATACGCTAGAGAGATAAGTATAGAATTCTGCAATGCATTGTCAAGTGTG
ATTCCATGTTATGCAGCAGGACAAATGCAAAAGCTTCAAGTGCTGACAGT
CAGTTCTTGTAATGGTCTGAAGGAGGTATTTGAAACTCAATTAAGGAGGA
GCAGCAACAAAAACAACGAGAAGAGTGGTTGTGATGAAGGAAATGGTGGA
ATTCCAAGAGTAAATAACAATGTTATTATGCTTTCTGGTCTGAAGATATT
GGAAATCAGCTTTTGTGGGGGTTTGGAACATATATTCACATTCTCTGCAC
TTGAAAGCCTGAGACAGCTCGAAGAGTTAACGATAATGAATTGCTGGTCA
ATGAAAGTGATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGCAAAC
AACAACAACAACGAAGGGGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
CTTCTTCTTCTTCTTCTTCTCCTCCTTCTTCTTCTAAGAAGGTTGTGGTC
TTTCCTTGTCTAAAGTCCATTGTATTGGTCAATCTACCAGAGCTGGTAGG
ATTCTTCTTGGGGATGAATGAGTTCCGGTTGCCTTCATTAGATGAACTTA
TCATCGAGAAATGCCCAAAAATGATGGTGTTTACAGCTGGTGGGTCCACA
GCTCCCCAACTCAAGTATATACACACAAGATTAGGCAAACATACTATTGA
TCAAGAATCTGGCCTTAACTTTCATCAGGTATATATGTTTCTTTAATTGG
CATCATCTAATTAAGAAAGATATCATTCCTGCCAAGTAAATTTACTTCAA
ACACATTCACACTGGTTTCAGTCTAAGTTTATGTTGTTCTAGGAAGGCCA
AAATGGGAAAGCAAGATAGGGAAAAATAGTGTATTTCAGTGGAAAGGGTA
TTTTAGGTATTTTCTGTCAAAAGTTGTTATTGCAGGCTTTTTAGTACCTG
GAATCGTGTGTGGGAGGAGCATTATTATTCTGATTTGCTTGTTTCTTTAT
CATTTTTTCTTAGCCTCTCGAACAGCTAGAAACCCTTTTAATCTTTTGAT
TTTAAATGACAAAATTTTTCCCTGTTACTCTATTTGATTGTTGTTCTTCA
TGGTTCTAAGTGAGTTATTGGCTCATCTGTTACTTCTTTTGATTGTTATT
TTCATAGCATGTTAGTCACTTGAATCAAGCTTTTTCATTTTCAACCAGGG
CAAAAGGTCAAAAGTAACCTACTTTATGAGATCAAAAACAGCAACCCATC
GGATAACTTTTAGTTGGAGTTAATAGTTACAATTACCATTGTGATTAATA
ATTATAATATCCTGTATTAATTCATAAAAATTGGTACAGCACATATATGA
CATTTCAAAGGTTTTTGTTTGACATATATATGCCTCTGGCGTTTTCTTTA
TTGGACTTGCAGACCTCATTCCAAAGTTTATACGGTGACACCTTGGGCCC
TGCTACTTCAGAAGGGACAACTTGGTCTTTTCATAACTTGATTGAATTAG
ATGTGAAATTTAATAAGGATGTTAAAAAGATTATTCCATCCAGTGAGTTG
CTGCAACTGCAAAAGCTGGAAAAGATAAATATAAACAGTTGTGTTGGGGT
AGAGGAGGTATTTGAAACTGCATTGGAAGCAGCAGGGAGAAATGGAAATA
GTGGAATTGGTTTTGATGAATCGTCACAAACAACTACCACTACTCTTGTC
AATCTTCCAAACCTTAGAGAAATGAACTTATGGGGTCTAGATTGTCTGAG
GTATATATGGAAGAGCAATCAGTGGACAGCATTTGAGTTTCCAAAACTAA
CAAGAGTTGAAATTAGTAATTGCAACAGTTTAGAACATGTATTTACTAGT
TCCATGGTTGGTAGTCTATCGCAACTCCAAGAGCTACATATAAGTCAGTG
CAAACTTATGGAGGAGGTGATTGTTAAGGATGCAGATGTTTCTGTAGAAG

AAGACAAAGAGAAAGAATCTGATGGCAAGATGAATAAGGAGATACTTGCG
TTACCTAGTCTAAAGTCCCTGAAATTAGAAAGCTTACCATCTCTTGAGGG
GTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATTATTGGATACTTTAA
GAATTGAGGAATGCCCAGCAATAACCACCTTCACCAAGGGAAATTCCGCT
ACTCCACAACTAAGAGAAATAGAAACAAGATTTGGCTCGGTTTATGCAGG
GGAAGACATCAAATCCTCTATTATAAAGATCAAACAACAGGTAAATCAGA
TCATTGTTGGTTTAATAATTCTTAAACTACATTTGAAAAGTTTCATGTAA
GTTTTTTATTATTGTCAAAAGCCGCAACCTATATTTTCAACTTTATATTT
ATGTACTTTATGCAGGATTTCAAAAAAGCCCAGGACTCTATTTAATGTGA
AGTAAATACTAGAAGAGGTAAATTCTATTTACATGTCTCCTGATTGCCTA
TTAATTAATGGCCTTTCAGTTCATGGTTTTTGGATGTATTCTTCATGATG
ACGTGAATGTTTAAATACCCCACTAGTTAATTGTTAGGTTGAATGTTGAT
GACCAAAGGACTATATGTCGGGAAGAATATTCAAGGAAAGAATTGTTCAT
CATATGAAGGGCATTAAATTAAGAAGAACATGGATGCTATGAAGATGTTG
GGAAAATATATGAATCAAATAACAAGCTACTCACTTATCTAAGTTTGTTG
GTTGAGGATGTTGATTTTAATATTTCAAATTCATTGGTATCATTATATGG
GTTTATCAGTAGTGTTAATGGGATAATGAGCAACTTAACCTTAAATTATG
CTGTTGGTAAATGTTGGACTCAAGTATGGAAAATTAGGAATAACTTGTGA
AAAATATATGCAAAAGTAGGATTGAGATTTTCAATGAAAAAAATTATGAA
ACTATACTACTATAGTATATAAATAAATTCAACTTACTGTTGGGTATATT
GGAAGCACATATCATGAAAGTAACTAGAAGCAGAATTTGTTCCCATCTTC
ATCTACTTATAGTTTCCATTTCTTACTTGTAAAAATCTGATTAAACTTTA
GAGTTATTTCTATTTTTTACCAACCAAAATTTTCATATAAAGGCCACAAG
T **(SEQ ID NO:107)**

RG2J deduced polypeptide sequence (SEQ ID NO:108)

MSDPTGIVGAIINPIAQTALVPLTDHVGYMISCRKYVRDMQMKMTELNTSRISAEEH
ISRNTRNHLQIPSQIKDWLDQVEGIRANVANFPIDVISCCSLRIRHKLGQKAFKITEQI
ESLTRQNSLIIWTDEPVPLGRVGSMIASTSAASSDHHDVFPSREQIFRKALEALEPVQ
KSHIIALWGMGGVGKTTMMKKLKEVVEQKKTCNIIVQVVIGEKTNPIAIQQAVADY
LSIELKENTKEARADKLRKRFEADGGKNKFLVILDDVWQFFDLEDIGLSPLPNKGV
NFKVLLTSRDSHVCTLMGAEANSILNIKVLKDVEGKSLFRQFAKNAGDDDLDPAFI
GIADSIASRCQGLPIAIKTIALSLKGRSKSAWDVALSRLENHKIGSEEVVREVFKISYD
NLQDEVTKSIFLLCALFPEDFDIPIEELVRYGWGLKLFIEAKTIREARNRLNNCTERL
RETNLLFGSHDFGCVKMHDVVRDFVLHMFSEVKHASIVNHGNMSEWPEKNDTSN
SCKRISLTCKGMSKFPKDINYPNLLILKLMHGDKSLCFPENFYGKMEKVQVISYDKL
MYPLLPSSLECSTNVRVLHLHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIG
NLKKLRLLDLTNCKGLRIDNGVLKNLVKLEELYMGVNRPYGQAVSLTDENCNEM
VEGSKKLLALEYELFKYNAQVKNISFENLKRFKISVGCSLHGSFSKSRHSYENTLKL
AIDKGELLESRMNGLFEKTEVLCLSVGDMYHLSDVKVKSSSFYNLRVLVVSECAEL
KHLFTLGVANTLSKLEHLKVYKCDNMEELIHTGGSEGDTITFPKLKLLYLHGLPNL
LGLCLNVNAIELPKLVQMKLYSIPGFTSIYPRNKLEASSLLKEEVVIPEELIVEKCGSI
EELFNIDLDCASVIGEEDNNSSLRNINVENSMKLREVWRIKGADNSRPLFRGFQVVE

KIIITRCKRFTNVFTPITTNFDLGALLEISVDCRGNDESDQSNQEQEQIEILSEKETLQE
ATDSISNVVFPSCLMHSFHNLQKLILNRVKGVEVVFEIESESPTSRELVTTHHNQQQP
VIFPNLQHLDLRGMDNMIRVWKCSNWNKFFTLPKQQSESPFHNLTTINIDFCRSIKY
LFSPLMAELLSNLKKVNIKWCYGIEEVVSNRDDEDEEMTTFTSTHTTTILFPHLDSL
TLSFLENLKCIGGGGAKDEGSNEISFNNTTATTAVLDQFELSEAGGVSWSLCQYAR
EISIEFCNALSSVIPCYAAGQMQKLQVLTVSSCNGLKEVFETQLRRSSNKNNEKSGC
DEGNGGIPRVNNNVIMLSGLKILEISFCGGLEHIFTFSALESLRQLEELTIMNCWSMK
VIVKKEEDEYGEQQTTTTTKGTSSSSSSSSSSSSSSSPPSSSKKVVVFPCLKSIVLVNLP
ELVGFFLGMNEFRLPSLDELIIEKCPKMMVFTAGGSTAPQLKYIHTRLGKHTIDQES
GLNFHQDIYMPLAFSLLDLQTSFQSLYGDTLGPATSEGTTWSFHNLIELDVKFNKD
VKKIIPSSELLQLQKLEKININSCVGVEEVFETALEAAGRNGNSGIGFDESSQTTTTTL
VNLPNLREMNLWGLDCLRYIWKSNQWTAFEFPKLTRVEISNCNSLEHVFTSSMVGS
LSQLQELHISQCKLMEEVIVKDADVSVEEDKEKESDGKMNKEILALPSLKSLKLESL
PSLEGFSLGKEDFSFPLLDTLRIEECPAITTFTKGNSATPQLREIETRFGSVYAGEDIKS
SIIKIKQQDFKKAQDSI.CEVNTR

RG2K polynucleotide sequence (SEQ ID NO:109) and (SEQ ID NO:110)

TGGGATTCCATATATAAAAACATATATTTTTATAAAGTGGGATTCCATTG
TTTATATAGATTTTTATTCACCAATAGACAATAGATTAAAAAAAGATATA
AAAACATGTCGGCTTTTGACTAAAAATATAGATTTTTATGAATAGAATAT
TCAATTTGCTTAACTCGTTTAAAAAAAAATGAAAAAGATGTCGATATAAAA
TCTCATATGGGCCTTCTTTACCATTCAAATAGTAAAATAGTAAAAGATAC
TTGTTTGGGGCATGAACTGACCATAGTCAAACCCATACAAAATCAAACGA
ATCCCACATGGATGATGACGATGGGGTCGCAGTAAATGTGTTTTGGTCCT
TTTTTTTCGAGAGAACAGAAGCTTCTGCTCTTCATCTTCTTTAGATTTTG
GGGATTTTCTGGTTTCAGGGGTTTGTGAGTGGAAACTAAATTGAAGCAAA
AAAGTATGGTATAATTGGTTGCTAGTGAAATTGATGCTTTCTATTACTAT
CATCTTTAAAATTGTCAAAACATTATGTATTAAATTATGAGATCGAAAGT
GGTCTATGGGCCAAAGGTAATACAAGCTTACTCAATGAAATGAATCTAGG
ATGCATCATGCATGTATTGGTTAGATTAAAGATTTTCATCAAATTTCCTT
TATCAAATTGTTGTATACCATGTTATGTAGGTGCTACCACAAGCCATAAC
ATCGAGCAATGGAGTGTATTACTGGCATCTTTAGCAACCCGTTTGCTCAG
TGTCTCATCGCTCCTGTGAAAGAACACCTTTGCCTTCTGATTTTCTATAC
ACAATATGTAGGGGATATGCTTACTGCAATGACGGAGTTGAATGCTGCAA
AAGACATTGTTGAAGAGCGGAAGAATCAAAACGTAGAAAAATGTTTTGAG
GTTCCAAACCATGTCAACCGTTGGTTGGAAGATGTTCAAACAATCAACAG
AAAAGTGGAACGTGTTCTTAACGATAATTGCAATTGGTTCAATCTATGTA
ATAGGTACATGCTCGCAGTGAAAGCCTTGGAGATAACTCAGGAGATCGAT
CATGCCATGAAACAACTCTCTCGGATAGAATGGACTGATGATTCAGTTCC
TTTGGGAAGAAATGATTCCACAAAGGCATCCACCTCTACACCATCAAGTG
ATTACAATGACTTCGAGTCAAGAGAACACACTTTTAGGAAAGCACTTGAA
GCACTTGGATCCAACCACACATCCCACATGGTAGCCTTATGGGGGATGGG

TGGAGTTGGGAAGACCACGATGATGAAGAGGCTGAAAAATATTATTAAAG
AAAAGAGGACGTTTCATTATATTGTTTGGTGGTTATAAAGGAAAATATG
GATCTCATTTCCATCCAGGATGCTGTAGCAGATTATCTGGATATGAAGCT
AACAGAAAGCAATGAATCAGAAAGAGCCGATAAACTTCGTGAAGGGTTTC
AGGCCAAATCAGATGGAGGTAAGAATAGGTTCCTCATAATACTGGATGAT
GTATGGCAATCTGTTAATATGGAAGATATTGGTTTAAGTCCTTTTCCGAA
TCAAGGTGTCGACTTCAAGGTCTTGTTGACCTCGGAAAACAAAGATGTTT
GTGCAAAAATGGGAGTTGAAGCTAATTTAATTTTCGACGTGAAATTCTTA
ACAGAAGAAGAAGCACAAAGTTTGTTTTATCAATTTGTAAAAGTTTCTGA
TACCCACCTTGATAAGATTGGAAAAGCTATTGTAAGAAACTGTGGTGGTC
TACCCATTGCCATCAAAACCATAGCCAATACTCTTAAAAATAGAAACAAG
GATGTATGGAAGGATGCACTTTCTCGTATAGAGCATCATGACATTGAGAC
AATTGCACATGTTGTTTTTCAAATGAGCTACGACAATCTCCAAAACGAAG
AAGCTCAATCCATTTTTTTGCTTTGTGGATTGTTTCCTGAAGACTTTGAT
ATTCCTACTGAGGAATTGGTGAGGTATGGATGGGGATTGAGAGTATTTAA
TGGAGTGTATACTATAGGAGAAGCAAGACACAGGTTGAACGCCTACATCG
AGCTGCTCAAGGATTCTAATTTATTGATTGAAAGTGATGATGTTCACTGC
ATCAAGATGCATGATTTAGTTCGTGCTTTTGTTTTGGATACGTTTAATAG
ATTCAAGCATTCTTTGATTGTTAACCATGGTAATGGTGGTATGTTAGGGT
GGCCTGAAAATGATATGAGTGCCTCATCTTGCAAAAGAATTTCATTAATA
TGCAAGGGCATGTCCGATTTTCCTAGAGACGTAAAGTTTCCAAATCTCTT
GATTTTGAAACTTATGCATGCAGATAAGTCTTTGAAGTTTCCTCAAGACT
TTTATGGAGAAATGAAGAAGCTTCAGGTTATATCATACGATCACATGAAG
TATCCCTTGCTTCCAACATCACCTCAATGCTCCACCAACCTTCGTGTGCT
TCATCTTCATCAATGCTCATTGATGTTTGATTGCTCTTCTATTGGAAATC
TGTTGAATCTGGAAGTGCTCAGCTTTGCTAATTCTGGTATTGAGTGGTTG
CCTTCCACAATCGGAAATTTGAAGGAGCTAAGGGTACTAGATTTGACAAA
TTGTGATGGTCTTCGTATAGATAATGGTGTCCTAAAGAAATTGGTGAAAC
TTGAAGAGCTTTATATGAGAGTTGGTGGTCGATATCAAAAGGCCATTAGC
TTCACTGATGAAAACTGCAATGAAATGGCAGAGCGTTCAAAAAATCTTTC
TGCATTAGAATTTGAGTTCTTCAAAAACAATGCTCAACCAAAGAATATGT
CATTTGAGAATCTTGAACGATTCAAGATCTCAGTGGGATGTTATTTTAAG
GGAGATTTCGGTAAGATCTTTCACTCTTTTGAAAACACGTTGCGGTTGGT
CACCAACAGAACTGAAGTTCTTGAATCTAGGCTTAATGAGTTGTTTGAGA
AAACAGATGTTCTTTATTTAAGTGTGGGAGATATGAATGATCTTGAAGAT
GTTGAGGTAAAGTTGGCACATCTTCCTAAATCCTCTTCCTTCCACAATTT
AAGAGTCCTTATCATTTCTGAGTGTATAGAGTTGAGATACCTTTTCACAC
TTGATGTTGCAAACACTTTGTCAAAGCTTGAGCATCTTCAAGTTTACGAA
TGCGATAATATGGAAGAAATCATACATACAGAGGGTAGAGGAGAAGTGAC
AATTACATTCCCAAAGCTGAAGTTTTTATCATTGTGTGGGCTACCAAATC
TGTTGGGTTTGTGTGGTAATGTGCACATAATTAATCTACCACAACTCACA
GAGTTGAAACTTAATGGCATTCCAGGTTTCACAAGCATATATCCTGAAAA
AGATGTTGAAACATCTAGTTTGTTGAATAAAGAGGTAAATGTGTTTTATG

TTAATACAATACAATCTTTTCAATTAACCGTTTCAAAATATATTGTATGA
TTTATTTTTGTTTGGATGGGGTTATTAATGGGTGATTATTTCTCAGGTTG
TAATTCCTAATTTGGAGAAACTTGATATTAGTTATATGAAGGATTTGAAA
GAGATATGGCCTTGTGAATTAGGGATGAGTCAGGAAGTTGATGTTTCTAC
GTTGAGAGTGATTAAAGTAAGCAGTTGTGATAATCTTGTGAATCTATTCC
CGTGCAATCCTATGCCATTGATACATCACCTTGAAGAGCTTCAAGTGATA
TTTTGTGGTTCCATTGAAGTGTTATTCAACATTGAGTTGGATTCTATTGG
TCAAATTGGAGAAGGCATCAACAATAGCAGCTTGAGAATCATCCAATTGC
AGAACTTAGGGAAGCTAAGTGAGGTGTGGAGGATAAAAGGTGCGGATAAC
TCTAGTCTTCTCATCAGTGGCTTTCAAGGTGTTGAAAGCATTATCGTTAA
CAAATGCAAGATGTTTAGAAATGTATTCACACCTACCACCACCAATTTTG
ATCTGGGGGCACTTATGGAGATTCGGATACAAGATTGTGGAGAAAAGAGG
AGAAACAACGAATTGGTAGAGAGTAGCCAAGAGCAAGAGCAGGTATGGCT
TTCAATTTCACTTTCTTACTTAATGAAGGATTAAGCTCCTGCTTTTTGAA
TAAAAAGTGGATGAATGACTAAATTCGGGAATGCCACCCGGAAAGTTATC
AACCATTTAGCTACACCATTTTTTGAACTAATGTTGCAATAAATGCATAA
TATAATTAAAAAATGGTCATTGATAAATGTAAACCAACCTTTTTTATTTA
TTAAAATGTCTACAATAAATGATTTTCTTTATTATATATCATTTTATAAC
AATAAGCTTAAAGATGTTTAAATAGCCAATGTCAGTTATAGATCGTAACT
AATTTTTTATTAACTAGTTTTAGTTAAGATATCACTCATTATTATTTTTA
TAGAAAAAAGACAAGATTGGCTAATCCTCATAAGAATTTGGAAGATTTAA
GCAAAATATAGAGCTTTTCCAAACATAGCCAATAGTTTCTTTTGCAGGTC
CCATCTACGAAATTATCAATAGATTTGCGATTTTTTTTTGGCACCCGGGA
AATTTCCATTAATTAAAAAAAAGTTCAAGCCATTTTGTAGTTGGCACCTG
CAAAATGGTAGTTTGCACCTGCGGAAATCACCTTTCACCATTTCGCATCT
ATGACTTGTGAAAATGTTAATTTGTGAAATGGTCATGTGCACCTCATGAG
AAATACGAAATGGTCAGTAATATGACTTTTTTATATAAATATGATGGTGG
CATATATTTATAGGAAAATATAGCTGCACGATATTAATTAATAGTGAAAT
TAGTTAACTGTATACGATAAGTATACAAAATTTATATGTATGAAGTATAC
TCAATTTAGGACGACTCGGGCAATGAAATCATCATTTAATAGGAGCAATG
AAATCATTTTCGAAAAATGTTTACAAATGAATAAAATATTAAATTAAACT
TAAAACATTTTGTTAGTAGTTTGAAATTTACAAACTGAAATTTGTTGTAT
TTATTAACATTTATAAATGTTGTACTATGATTTTTTCCTTGTTTGCAAAT
ATTCCTTAAAAATCCACCTAAAATCAAAATAATTAATCTTTTTCAAGTTG
AAAAATGAAAATCGTATGATATAACCGTGTATGGATGTGGAATTATATAT
CAGTTACTAATTACATTTTTTGTTGGGATATATGTGCGCAGATTGATATT
GCAATCCCATTCACTCTCACACACTCTTTCCAAAACCTCCGTAAACTTGC
TTTGGAAAAGTATGAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTCCAA
CAAGTAGAGAATTGATAACAATTCACCATAATCAACAACCACTACTTCCC
AACCTTGAGTTATTGGATATAAGTTTTATGGACAGCATGAGTCATGTATG
GAAGTGCAACTGGAATAAATTCTTCATTCTTCAAAAACAACAGTCAGAAT
CCCCATTCTGTAATCTCACAACCATACATATTCAATATTGCCAAAGCATT
AAGTACTTGTTTTCAACTCTCATGGCAAAACTTCTTTCCAACCTAAAGAA

GGTCGAGGTAAGAGAGTGTCATGGTATTGAAGAAGTTGTTTCGAACAGAG
ATGATGAAGATGAGGAAAAGACTACATTTACATCTACATCTTCTGAAAAA
AGCACTAATTTGTTCCCTCGTCTTGAATCTCTCGCTCTTTATCAACTTCC
AAATCTCAAGTGTATTGGTGGTGGTGGTTCTGCCAACAGTGGGAACAATG
AAATATCTCTTGATAATTCCACTACTACTACTTCTTTTGTTGATCAATCT
AAGGTATGTTTTTTTTTTTNGTTNCCCTT **(SEQ ID NO:109)**

Sequence gap

CCTCCCTAATAATACATGTTATGCACACTATACTAACATATTAGACACGT
AAAGGATAAATGCTATGCCTCATATAATACGTTATATTTATAATCTTTAA
ACAATCAAATTTATTAAACAAATAACTAAGTGTGAGCAAAGGCAGGTACC
CGACTAAATTGCCCAAAACCAGTCTGGTGGTTCGTGGAATGTTGGGCCAG
GTCGTTAAAACGTCTACACACCGGTTCTTTAAATCACAGATCCGCTTCTC
ATACTGTGAACCCGGTTTTAATTTTAAAAGAAAATTTCATTATAAAGTAA
ATGACTTAAACCATTACAAACAACAAAAATTTACCATTACAATGTTGGAC
TATCATTATTTGCAACATAAAACTGAAAATACACATATTTCCTTCTGATA
TCAGCATGAGTGGCTGGTTGGCTAACCCAAAAATCCATGCATTGTAGATG
TGTGTTACAACACATAGTATCAATGAAAGGCATATTTTTAGGCTAGAATT
TAACAATCTGTAATAATATTCCCTAAAACTAATATCATCATCAACCAACT
AATATAAAACCATTGGGTTCGTCATTTTAGGTACAAAACATAGATTTTTC
TAAGCTTGTTGTATTTAAACATATGCTTTCTAAACTTAATTGATTTTGCA
TTCCAAAATTTTAGGTTGTAAAGTGGTATGTCATTTGTTGTCTTTTCAAC
ATTAATTGTACAAAAACCAAAACTACATAATTGATGTAGATATCATAACA
ATTGTGTTATTTAGTATATAAAAACTAAATTTTGAATTGAATTTCTTATA
CAAAAGTTGTGTCTATGTATACATGTTTATGTAGGTAATAGACAATTAGT
CTCTGTTAAGTATATGGAGTTTAATTTTTAGACTAATTTTTCATGTGTTG
CAGTTTTATCAGGCAGGTGGCGTTTTTTGGACGTTATGCCAATACTCCAG
AGAGATAAATATAAGGGAGTGTTATGCATTGTCAAGTGTAATTCCATGTT
ATGCAGCAGGACAGATGCAAAATGTTCAAGTGCTGAATATATACAGGTGC
AACTCAATGAAGGAGTTATTTGAAACTCAAGGGATGAACAACAACAATGG
TGACAGTGGTTGTGATGAAGGAAATGGTTGTATACCAGCAATTCCAAGAC
TAAATAACGTTATTATGCTACCCAATCTAAAGATATTGAAGATTGAAGAT
TGTGGTCATCTGGAACATGTATTCACATTCTCTGCACTTGGAAGCCTGAG
ACAGCTCGAAGAGTTAACGATAGAGAAATGCAAGGCAATGAAAGTGATAG
TGAAGGAAGAAGATGAATATGGAGAGCAAACAACAAAGGCATCTTCGAAG
GAGGTTGTGGTCTTTCCTCGTCTCAAGTCCATTGAACTGGAAAATCTACA
AGAGCTCATGGGTTTCTACTTAGGGAAGAATGAGATTCAGTGGCCTTCAT
TGGATAAGGTTATGATCAAGAATTGCCCAGAAATGATGGTGTTTGCACCT
GGTGAGTCCACAGTTCCCAAGCGCAAGTATATAAATACAAGCTTTGGCAT
ATATGGGATGGAGGAGGTACTTGAAACTCAAGGGATGAACAACAATAATG
ATGACAATTGTTGTGATGATGGAAATGGTGGAATTCCAAGACTAAATAAC
GTTATTATGTTTCCAAATATAAAGATATTGCAAATCAGCAATTGTGGCAG
TTTGGAACATATATTCACATTCTCTGCACTTGAAAGCCTGATGCAGCTCA
AAGAGTTAACAATAGCGGATTGCAAGGCAATGAAAGTGATTGTGAAGGAG

GAATATGATGTAGAGCAAACAAGGGTATTGAAGGCTGTGGTATTTTCTTG
TCTAAAGTCCATTACACTATGCCATCTACCAGAGTTGGTGGGTTTCTTCT
TGGGGAAGAATGAGTTCTGGTGGCCTTCATTGGATAAGGTTACCATCATT
GATTGCCCACAAATGATGGGGTTCACACCTGGTGGGTCAACAACTTCCCA
CCTCAAGTACATACACTCAAGCTTAGGCAAACATACTCTTGAATGTGGCC
TTAATTTCAAGTCACAACTACTGCATATCATCAGGTATAATTATTATTCT
TTNACACCATCTAATTATGGAATCATGACGCTAATTACAGTATTAAACAC
(SEQ ID NO:110)

RG2K deduced polypeptide sequence (SEQ ID NO:111)

MECITGIFSNPFAQCLIAPVKEHLCLLIFYTQYVGDMLTAMTELNAAKDIVEERK
NQNVEKCFEVPNHVNRWLEDVQTINRKVERVLNDNCNWFNLCNRYMLAVKAL
EITQEIDHAMKQLSRIEWTDDSVPLGRNDSTKASTSTPSSDYNDFESREHTFRKAL
EALGSNHTSHMVALWGMGGVGKTTMMKRLKNIIKEKRTFHYIVLVVIKENMDL
ISIQDAVADYLDMKLTESNESERADKLREGFQAKSDGGKNRFLIILDDVWQSVN
MEDIGLSPFPNQGVDFKVLLTSENKDVCAKMGVEANLIFDVKFLTEEEAQSLFY
QFVKVSDTHLDKIGKAIVRNCGGLPIAIKTIANTLKNRNKDVWKDALSRIEHHD
IETIAHVVFQMSYDNLQNEEAQSIFLLCGLFPEDFDIPTEELVRYGWGLRVFNGV
YTIGEARHRLNAYIELLKDSNLLIESDDVHCIKMHDLVRAFVLDTFNRFKHSLIV
NHGNGGMLGWPENDMSASSCKRISLICKGMSDFPRDVKFPNLLILKLMHADKS
LKFPQDFYGEMKKLQVISYDHMKYPLLPTSPQCSTNLRVLHLHQCSLMFDCSSI
GNLLNLEVLSFANSGIEWLPSTIGNLKELRVLDLTNCDGLRIDNGVLKKLVKLEELY
MRVGGRYQKAISFTDENCNEMAERSKNLSALEFEFFKNNAQPKNMSFENLERFKIS
VGCYFKGDFGKIFHSFENTLRLVTNRTEVLESRLNELFEKTDVLYLSVGDMNDLED
VEVKLAHLPKSSSFHNLRVLIISECIELRYLFTLDVANTLSKLEHLQVYECDNMEEII
HTEGRGEVTITFPKLKFLSLCGLPNLLGLCGNVHIINLPQLTELKLNGIPGFTSIYPEK
DVETSSLLNKEVVIPNLEKLDISYMKDLKEIWPCELGMSQEVDVSTLRVIKVSSCDN
LVNLFPCNPMPLIHHLEELQVIFCGSIEVLFNIELDSIGQIGEGINNSSLRIIQLQNLGK
LSEVWRIKGADNSSLLISGFQGVESIIVNKCKMFRNVFTPTTTNFDLGALMEIRIQDC
GEKRRNNELVESSQEQEQ

RG2L polynucleotide sequence (SEQ ID NO:112)

GGAAGACACAATGATGCAAAGACTGAAGAAGGTTGCCAAAGAAAATAGAA
TGTTCAGTTACATGGTCGAGGCAGTTATAGGGGAAAAGACAGACCCAATT
GCTATTCAACAAGCTGTAGCCGATTACCTTCGTATACAGTTCAAAGAAAG
CACTAAACCAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAGGCCCACT
CTGNAGACGGTAAGAATAAGTTCCTCGTAATATTTGATGACGTCTGGCAG
TCCGTTGATCTGGAAGATATTGGNTTAAGTCCTTTTCCAAATCAAGGTGT
CGACTTCAAGGTCTTGTTGACTTCACGAGACGAACACGTTTGCACAATGA
TGGGGGTTGAAGCTAATTCAGTTATTAATGTGGGACTTCTAACTGAAGTA
GAAGCACAAAGTCTGTTCCAGCAATTTGTAGAAACTTTTGAGCCCGAGCT
CTGTAAGATAGGAGAAGTTATCGTAAGAAAGTGTTGCGGTCTACCTATTG

CCATCAAAACCATGGCGTGTACTCTAAGAAATAAAAGAAAGGATGCATGG
AAGGATGCACTTTCACGTATAGAGCACTATGACATTCGTAGTGTTGCGCC
TAAAGTCTTTGAAACAAGCTATCACAATCTCCAAGACAGGGAGACTAAAT
CCGTGTTTTTGATGTGTGGTTTGTTTCCTGAAGACTTCAATATTCCTACC
GAGGAGTTGATGAGGTATGGATGGGGCTTAAAGCTATTTGACAGAGTTTA
TACAATTAGAGAAGCAAGAACCAGGCTCAACACCTGCATTGAGCGACTTG
TGCAGACAAATTTGTTAATTGAAAGTGATGATGTTGGGTGTGTCAAGATG
CATGATCTGGTGCGTGCTTTTGTTTTGGGTATGTATTCTGAAGTCGAGCA
TGCTTCAATTGTCAACCATGGTAATATGCATGGGTGGACTAAAAATGATA
TGAACGACTCTTGCAAAACAGTTTCTTTAACATGCGAGAGTGTGTCTGAG
TTTCCAGGAGACCTCAAGTTTCCAAACCTAAAGCTTTTGAAACTTATGCA
TGGAGATAAGATGCTAAGGTTTTCTCAAGACTTTTATGAAGGAATGGAAA
AGCTCCAGGTAATATCATACCATAAAATGAAGTATCCATTGCTTCCCTCG
TCACCTCAATGCTCCACCAACCTTCGAGTGCTTCATCTTCATCGGTGTTC
ATTACGGATGCTTGATTGCTCTTGTATCGGAAATTTGACGAATCTGGAAG
TGTTGAGCTTCGCTAATTCTGGCATTGAACGGATACCTTCAGCAATCGGA
AATTTGAAGAAGCTTAGGCAACTTGATCTGAGAGGTCGTTATGGTCTTTG
TATAGAACAGGGTGTCTTGAAAAATTTGGTCGAACTTGAAGAACTTTATA
TTGGAAATGCATCTGCGTTTAGAGATTATAACTGCAATGAGATGGCAG

RG2L deduced polypeptide sequence (SEQ ID NO:113)

EDTMMQRLKKVAKENRMFSYMVEAVIGEKTDPIAIQQAVADYLRIQFKESTKPAR
ADKLREWFKAHS?DGKNKFLVIFDDVWQSVDLEDIGLSPFPNQGVDFKVLLTSRDE
HVCTMMGVEANSVINVGLLTEVEAQSLFQQFVETFEPELCKIGEVIVRKCCGLPIAI
KTMACTLRNKRKDAWKDALSRIEHYDIRSVAPKVFETSYHNLQDRETKSVFLMCG
LFPEDFNIPTEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQTNLLIESDDVGC
VKMHDLVRAFVLGMYSEVEHASIVNHGNMHGWTKNDMNDSCKTVSLTCESVSEF
PGDLKFPNLKLLKLMHGDKMLRFSQDFYEGMEKLQVISYHKMKYPLLPSSPQCST
NLRVLHLHRCSLRMLDCSCIGNLTNLEVLSFANSGIERIPSAIGNLKKLRQLDLRGR
YGLCIEQGVLKNLVELEELYIGNASAFRDYNCNEMA

RG2M polynucleotide sequence (SEQ ID NO:114)

GGGGAAGACACAATAGATGCAAAGGCTGAAGAAGTTGCCAAAGAAAAGAG
AATGTTCAGTTATATCATTGAGGCGGTTATAGGGGAAAAGACAGACCCCA
TTTCCATTCAGGAAGCTATATCATATTACCTTGGTGTAGAGCTCAATGCA
AATACTAAGTCAGTAAGAGCTGATATGCTTCGTCAAGGGTTCAAGGCCAA
ATCTGATGTAGGTAAGGATAAATTCTTAATAATACTCGACGATGTATGGC
AGTCTGTTGATTTGGAAGATATTGGATTAAGTCCATTTCCAAATCAAGGT
GTTAACTTCAAGGTCCTGTTAACATCACGAGACCGACATATTTGCACTGT
GATGGGGGTTGAAGGTCATTCGATTTTTAATGTGGGACTTCTCACAGAAG
CAGAATCAAAAAGATTGTTCTGGCAGTTTGTAGAAGGTTCTGATCCTGAG
CTCCATAAGATAGGAGAAGATATTGTAAGTAAGTGTTGTGGTCTACCCAT

TGCCATTAAAACCATGGCATGTACACTTAGAGATAAAAGTACGGATGCAT
GGAAGGATGCACTGTCTCGTTTAGAGCATCATGACATTGAAAATGTTGCC
TCTAAAGTTTTTAGAGCGAGCTATGACCATCTCCAAGACGAGGAGACTAA
ATCCACTTTTTTTCTATGTGGATTGTTTCCAGAAGATTCCAATATTCCTA
TGGAGGAGTTGGTGAGGTATGGGTGGGGATTGAAATTATTTAAAAAAGTG
TATACCATAAGAGAAGCAAGAACTAGGCTCAACACTTGCATTGAGCGGCT
CATCTATACCAATTTGTTGATAAAAGTTGATGATGTTCAGTGCATCAAGA
TGCATGATCTCATCCGTTCTTTTGTTTTGGATATGTTTTCTAAAGTTGAG
CATGCTTCGATTGTCAACCATGGTAATACGCTAGAGTGGCCTGCAGATNA
TNTGCACGACTCTTGTAAAGGGCTTTCATTAACATGCAAGGGTANATGTG
AGTTTTGTGGAGACCTNAANTTTCCAACCCTAATGATTTTAAAACTTATG
CATGGAGATAAATCGCTAAGGTTT

RG2M deduced polypeptide sequence (SEQ ID NO:115)

GEDTIDAKAEEVAKEKRMFSYIIEAVIGEKTDPISIQEAISYYLGVELNANTKSVRAD
MLRQGFKAKSDVGKDKFLIILDDVWQSVDLEDIGLSPFPNQGVNFKVLLTSRDRHI
CTVMGVEGHSIFNVGLLTEAESKRLFWQFVEGSDPELHKIGEDIVSKCCGLPIAIKT
MACTLRDKSTDAWKDALSRLEHHDIENVASKVFRASYDHLQDEETKSTFFLCGLFP
EDSNIPMEELVRYGWGLKLFKKVYTIREARTRLNTCIERLIYTNLLIKVDDVQCIKM
HDLIRSFVLDMFSKVEHASIVNHGNTLEWPAD??HDSCKGLSLTCKG?CEFCGDL?F
PTLMILKLMHGDKSLRF

RG2N polynucleotide sequence (SEQ ID NO:116)

AGGTAAAATCCATAACCCTAAATGTTGGTACGCTCATATATCAAATTGCG
TGTTTTGTTGAATGAAAAAAGCATGCTCAAAAAACCAGTGTAAGGCACGG
TATATGACATATTTATAGTTACTGATAACAAATTATGATAATTTTGGGTT
TACRTAAGTTAGGATTCGTACTTCAACCAAATGTAATAGTTTTTGTGAGT
CTATCTATGTATTTGGGGAATCACATTAGCAACGGGATTGTACTAGTAAT
TCGAAAAAGTCTTTTAAATAATTTTTCTGTTTATAATTTATGAATAGTTT
TAGCGACATCTAATATTAAATAGAATGTATCTGATATTGAATTAATGTCC
TTAATGTGAACATAGACCTTTTCCATTTACTAATGCCTAATTATTAGTTT
CTAATCAATAAATTTTAATTTCTGTTTTATGCTTCTAAGACAATAAAAAT
CCATGATTTACCTTTAAATATTAACAAAAATGACCATAAATAAATAAAAA
ATTAGGATACCAAACCCCCCCGCCATGCCCAATGTCTAAATATTCTTGAT
GCTTTTGCTTTTCCCTCTTTTCCTTGTTAGTCTATTATTCTGGAGAGTTT
GAGAGAGTTTCATACAAGAAAATTTCAAGAAGAAAGCAAAGGTCCAGGTA
TTCTCTTTTCTTAATTATGTATTAACTTACAAGCATTTTTTACACGATCC
ATGGTTTTTTGTGTATGTTTTTCAAATTGAAACTAGATTGGGACTTTTGC
CCTTGATGATTCATAAGATATTGCATGGAGTTGAGATTGTGTAAGAAAAG
TGGTGAATAGAAAGAGCAAGTGAATCCAGATATAGTATTGGTAATATATG
ATGATGAGATAGAGATATGTTAAAACTGGCTAGAAAATTGTTTTAATTTG
AAATTTAGGTKGTTGAATTTGAAAGATACCAAGCTAATAACTAATTAGTT

ATGCTAAWTAGTTATAAAGAACAACAAACTCTTAGTTTTTTTTTCATGA
TTTTCAACCTCTTTGTACCAAACTAAATTATAGCAAAATTGAATATCATT
CTCTGCAATCAATCTTAACTTTTGTTATTATCATCATGTCTAAAATTGCC
ACAAGTTTATTTTCAAAGTCATATTGGATTATGAAAGGACTATTTTTACC
AATTACATCTTTACTTTATGGGCCAAAGCTAATACAATCCGACTAAACTA
AAGGAATATGGGATGCATATAGTTTGCTTCCCGATTATAGATTTCTATCT
AATTTGTCTATTGTACTAATTTAGGTGCCACCACAAGTAAATTTGTTAAA
TGGATATCGTTAATGCCATTCTTAAACCAGTTGTCGAGACTCTCATGGTA
CCCGTTAAGAAACACATAGGGTACCTCATTTCCTGCAGGCAATATATGAG
GGAAATGGGTATCAAAATGAGGGGATTGAATGCTACTAGACTTGGTGTCG
AAGAGCATGTGAACCGGAACATAAGCAACCAGCTTGAGGTTCCAGCCCAA
GGCAGGGGTTGGTATGAAGAAGTAGGAAAGATCAATGCAAAAGTGGAAAA
TTTTCCTAGCGATGTTGGCAGTTGTTTCAATCTTAAGGTTAGACACGGGG
TCGGAAAGAGAGCCTCCAAGATAATTGAGGACATCGACAGTGTCATGAGA
GAACACTCTATCATCATCTGGAATGATCATTCCATTCTTCTAGGAAGAAT
TGATTCCACGAAAGCATCCACCTCAATACCATCAACCGATCATCATGATG
AGTTCCAGTCAAGAGAGCAAACTTTCACAGAAGCACTAAACGCACTCGAT
CCTAACCACAAATCCCACATGATAGCCTTATGGGGAATGGGCGGAGTGGG
GAAGACGACAATGATGCATCGGCTGAAAAAGGTTGTGAAAGAAAAGAAAA
TGTTTAATTTTATTGTTGAGGCGGTTGTAGGGGAAAAAACAGACCCCATT
GCTATTCAATCAGCTGTGGCAGATTACCTAGGTATAGAGCTCAATGAAAA
AACTAAACCAGCAAGAACTGAGAAGCTTCGTAAATGGTTTGTGGACAATT
CTGCTGGTAAGAAGATCCTAGTCATACTCGACGATGTATGGCAGTTTGTA
GATCTGAATGATATTGGTTTAAGTCCTTTACCAAATCAAGGTGTCGACTT
CAAGGTGTTGTTGACATCACGAGACAAAGATGTTTGCACTGAGATGGGAG
CTGAAGTTAATTCAACTTTTAATGTGAAAATGTTAATAGAAACAGAAGCA
CAAAGTTTATTCCACCAATTTGTAGAAATTTCGGATGATGTTGATCGTGA
GCTCCATAATATAGGAGTGAATATTGTAAGGAAGTGTGGCGGTCTACCCA
TTGTCATCAAAACCATGGCGTGTACTCTTAGAGGAAAAAGCAAGGATGCA
TGGAAGAATGCACTTCTTCGTTTAGTGAACTACAACATTGAAAATATAGT
GAATGGAGTTTTTAAAATGAGTTACGACAATCTCCAAGATGAGGAGACTA
AATCCACCTTTTTGCTTTGTGGAATGTTTCCCGAAGACTTTAATATTCCT
ACCGAGGAGTTGGTGAGGTATGGATGGGGGTTGAAATTATTTAAAAAAGT
GTATACTATAGGAGAAGCAAGAATCAGGCTCAACACATGCATTGAGCGGC
TCATTCATACAAATTTGTTGATTGAAGTTGATGATGTTAGGTGCATCAAG
ATGCATGATCTTGTCCGTGCTTTTGTTTTGGATATGTATTCTAAAGTCGA
GCATGCTTCCATTGTCAACCATGGTAATACACTAGAGTGGCATGTGGATA
ATATGCACAACTCTTGTAAAAGACTTTCATTAACATGCAAGGGTATGTCT
AAGTTTCCTACAGACCTCAAGTTTCCAAACCTCTCGATTTTGAAACTTAT
GCATGAAGATATATCATTGAGGTTTCCCAAAAACTTTTATGAAGAAATGG
AGAAGCTTGAGGTTATATCCTATGATAAAATGAAATATCCATTGCTTCCC
TCATCACCGCAATGCTCCGTCAACCTTTGCGTGTTTCATCTCCATAAATG
CTCGTTAGTGATGTTTGACTGCTCTTGTATTGGAAATCTGTCGAATCTAG

AAGTGCTTAGCTTTGCTGATTCTGCCATTGACCTGTTGCCTTCCACAATC
GGAATTTTGAAGAAGCTAAGGCTACTGGATTTGACAAATTGTTATGGTCT
TTGTATAGCTAATGGTGTCTTTAAAAAATTGGTCAAACTTGAAGAGCTCT
ATATGACAGTGGTTAATGGAGGAGTTCGAAAGGCGATCAGCCTCACTGAG
GATAACTGCAATGAGATGGCAGAACGTTCAAAAGACCTTTCTGCATTAGA
ACTTGAGTTCTTTGAAAACAATGCTCAGCCAAAGAATATGTCATTTGAGA
AGCTACAACGATTCCAGATCTCAGTGGGGTGCTATTTATATGGAGCTTCC
ATAAAGAGCAGGCACTCGTATGAAAACACATTGAAGTTGGTTATTGACAA
AGGTGAATTATTTGAATCTTGAATGAACGGCCTGTTTAAGAAAACAGAGG
TGTTATGTTTAAGTGTGGGAGATATGAATGATCTTGAAGATRTTGAGGTT
AAGTCATCCTCACAACYTCTTCAATCTTCTTCGTTCAACAATTTAAGAGT
CCTTGTCGTTTCAAAGTGTGCAGAGTTGAAACACTTCTTCACACCTGGTG
TTGCAAACACTTTAAAAAAGCTTGAGCATCTTGAAGTTTACAAATGTGAT
AATATGGAAGAACTCATACGTAGCAGGGGTAGTGAAGAAGAGACGATTAC
ATTCCCCAAGCTGAAGTTTTTATCTTTGTGTGGGCTACCAAAGCTATCGG
GTTTGTGCGATAATGTCAAAATAATTGAGCTACCACAACTCATGGAGTTG
GAACTTGACGACATTCCAGGTTTCACAAGCATATATCCCATGAAAAAGTT
TGAAACATTTAGTTTGTTGAAGGAAGAGGTAAATATAAATTTTTAATGCT
AATACATTACAAAGGATCTTTTCAGTTAAATCTTTCAAAATATATTGTAA
TTTGATTGTATGGGGTATTATTGTTGGATGGGACTATTAATAAATGATTA
TCTTGCAGGTTCTGATTCCTAAGTTAGAGAAACTGCATGTTAGTAGTATG
TGGAATCTGAAGGAGATATGGCCTTGCGAATTTAATATGAGTGAGGAAGT
TAAGTTCAGAGAGATTAAAGTGAGTAACTGTGATAAGCTTGTGAATTTGT
TTCCGCACAAGCCCATATCTCTGCTGCGTCATCTTGAAGAGCTTAAAGTC
AAGAATTGTGGTTCCATTGAATCGTTATTCAACATCCATTTGGATTGTGC
TGGTGCAACTGGAGATGAATACAACAACAGTGGTGTAAGAATTATTAAAG
TGATCAGTTGTGATAAGCTTGTGAATCTCTTTCCACACAATCCCATGTCT
ATACTGCATCATCTTGAAGAGCTTGAAGTCGAGAATTGTGGTTCCATTGA
ATCGTTATTCAACATTGACTTGGATTGTGCTGGTGCAATTGGGCAAGAAG
ACAACAGAAGCAGCTTAAGAAACATCAAAGTGGAGAATTTAGGGAAGCTA
AGAGAGGTGTGGAGGATAAAAGGTGGAGATAACTCTCGTCCCCTTGTTCA
TGGCTTTCAATCTGTTGAAAGCATAAGGGTTACAAAATGTAAGAGGTTTA
GAAATGTATTCACACCTACCACCACAAATTTTAATCTGGGGGCACTTTTG
GAGATTTCAATAGATGACTGCGGAGAAAACAGGGAAAATGACGAATCGGA
AGAGAGTAGCCATGAGCAAGAGCAGGTAAGGATTTCAATTTCACTTTCKT
ACTTAATTAATGATTAAGCTCCTGCTTTTRAATAAAAAAGGGACAAACC
ATTTCATGACTTAATGTAGCAATACAAGTCATGTATAAGAGTGACCAACT
CTTTTTTATTTATAAAATGACTACAAAATATTTTTTTTCATTAGAGATCA
TGTATAAATGTGACTAATTTTTCATCACCTAACTTTAGTTGATAAATCTT
TATAAATGTCACTAGTTACTTTTCAGTAAAATAACAAATTTAATAAATTA
TCAACAAAAAGCATCAACTAAAAAAATCCCACAACCCGTAATAATTTAAA
ATAAAAGGATTTAACATCTAATACGAACAATTTTTTTTCTAAACATGATT
TGGACCAAATATCACCAGCAACTCAAGTTTGGAATCGATTCAGCTTAAAA

CTTGACCARCATAATTAGATAGATGAGAGTTGAAGCTAAAGTGCCTATAT
AAGTTCGTTTCATCTTTTTCTTGATCTTGATAGCAAGTTGAATSATTTT
CTTCTTCAAAATTGATAAAAATCTACATTATAAAGAGACTAGCTTGAAAA
AAAATGGTCTAGGTGGGTCTTGGGTCTGGTAGATGAAGATGGAAGGGAGA
GTAGATTTCAAAGACACAAACACATCTTCATTTTATTTATTTATTTATTA
TTATTATTTTTGATATCTTGCTCATATTTGTTACAGATATGTGAGGTCT
ATTAATCTTTTTAAATATATAAAAAATAAATACATAAATGAGAAAATTAA
ATAAAGAATAAATTAATAAGGGCACAATAGTCTTTTTTGGTAAGACAAGG
ACCAAAAGCGCAACAAAAGTAAACAGTAGGGACCATCCGATTTAAAAAAT
TAATTAGGGACCAAAAACATAAATTCCCCCAAACCATAGGGACCATTCGT
GTAATTTACTCTTGCTTTTCGTTTTGTTCATATTTGGGTAACTATTTTTT
TTGTACATATCTAGGTAACGAACTTGTTGAAAGTGTTCACATCTACGATG
TGACCTACTACAACCGATCATAATGGTCATATATGAACACTTCCAACAAG
TTTGTTATCTAGGTGTGTACAAAAAAACGATAGTTACCATGATGTGAACA
TACCAAAAAATTAATTACCTTAGCAAGTTATTTTCCCATTTAGGTTGTAT
GGAAACAGTTCCGTGAGACCGTGACTTGGATGGTAGATAAATTTAGTAAA
CTTAACCCTTCAATTAACCTACCTTTTTCTTATTAACTCAATTTCAAGCT
AAATTCTGATTCTTGTTTGAAAGTAAGTTGCATCTTTATGTTTGTATTAT
CTTGTTGCATAGGATCCTTAGCATCTTTTAATAATTTATTTGAAGGTGAA
AGATCCAACTATTTTTAATCTGTTGGCATTTTCCATCATTTGCAACTGTT
TCTTGAAAAAAA::TACCTAAAATCAAAATAACCATTTTCATATCCAAAA
TTATAAGAGAGAATTGTTAACGGACATGGAATCATAAATCATTAACACAG
TTCAGTACACAGGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTC
TATCAGAGAAAGAGACATTACAAGAAGCCACTGGCAGTATTTCAAATATT
GTATTCCCATCCTGTCTCATGCACTCTTTTCATAACCTCCATAAACTTAA
CTTGAACAGAGTTGAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTGAGA
GTCCAACAAGTAGAGAATTGGTAACAACTCACCATAACCAACAACAACCT
ATTATACTTCCCAACCTCCAGGAATTGATTCTATGGAATATGGACAACAT
GAGTCATGTGTGGAAGTGCGGCAACTGGAATAAATTCTTCACTCTTCCAA
AAGAACAATCAGAATCCCCATTCCACAACCTCAGTAACATACATATTTAT
GAATGCAAAAGCATTAAGTACTTGTTTTCACCTCTCATGGCAGAACTTCT
TTCCAACCTAAAGCATATCGAGATAAGAGAGTGTGATGGTATTGAAGAAG
TTGTTTCAAAAAGAGATGGTGAGGATGAAGACATGACTACATCTAC::::
:::GCACACAACCACCACTTTTTCCCTCATCTTGATTCTCTCACTCTAAA
GCAACTGAAGAATCTGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGG
GGAGCAATGAAATATCTTTCAATAATACCACTGCAACTACTGCTGTTCTT
GATCAATTTGAGGTATGCTTTGTACATATTCAATTATTTATTTAATTTCC
TTGTTAATTTCCTTTTTTCTTTGCAATATTCTATGAAAAAAATCACCAAA
TCACAAATAAGAGATTTAAACTTTTATTTCACACCCATGCGGACTCAAGA
ATGGGATTTGGAGGCATATAAAGTTACATTCATTTGAACAAGTATTACCA
TTTATTTGTTATTTATCATTTTCATATCATTTACTGATAACATTTCTTTT
TTACTTTTCTAATTAGAAAAGGTCCACATGTCTAATTAGGTTTTCCATTC
TATGTGAATCCTCTATTCTGTCTGTAATCAAGCATCTTAGATTATTTATC

CATTTTCATAATTGTGTTTATATTGACAGTTTTTTCTTTTTATAGTTGT
AATTGCAACCTGTCATATWTTMWWKKCWWWATKYWMWWARTAATACATTT
TATACCCWCTATACTAAGATA

RG2N deduced polypeptide sequence (SEQ ID NO:117)

LGKTTMMHRLKKVVKEKKMFNFIVEAVVGEKTDPIAIQSAVADYLGIELNEKTKPA
RTEKLRKWFVDNSAGKKILVILDDVWQFVDLNDIGLSPLPNQGVDFKVLLTSRDKD
VCTEMGAEVNSTFNVKMLIETEAQSLFHQFVEISDDVDRELHNIGVNIVRKCGGLPI
VIKTMACTLRGKSKDAWKNALLRLVNYNIENIVNGVFKMSYDNLQDEETKSTFLL
CGMFPEDFNIPTEELVRYGWGLKLFKKVYTIGEARIRLNTCIERLIHTNLLIEVDDVR
CIKMHDLVRAFVLDMYSKVEHASIVNHGNTLEWHVDNMHNSCKRLSLTCKGMSK
FPTDLKFPNLSILKLMHEDISLRFPKNFYEEMEKLEVISYDKMKYPLLPSSPQCSVNL
CVFHLHKCSLVMFDCSCIGNLSNLEVLSFADSAIDLLPSTIGILKKLRLLDLTNCYGL
CIANGVFKKLVKLEELYMTVVNGGVRKAISL

RG2O polynucleotide sequence (SEQ ID NO:118)

TTGTAAAACGACGGCCAGTCGAATCGTAACCGTTCGTACGAGAATCGCTG
TCCTCTCCTTCATTTGAATCATGATATTTGAATATCGATACTTTTGACTG
TAGCTTTTGGGTCGATTTTTAGCAAGATACATAACTGGCCAAACCCATT
GGCTATTTAGCCCAAAATATGAAATGGACTGGATTGTTTTTTCCTTTC
TAACACGCACACATCTGGCGATCAGTATCACTCCATTATGAAGACCTAGT
CAAATTCATTAACGTTCAGTCGTTCCTTCAAAGTTTCAAAGTTCCAACTT
CCAACTTCCCTCTTTTTTTTTCTTTCCTCGATTCTGATTTGAATCCGAT
TCTGCGACGAAGGAGAGCTTGGTCAGAGGGCTGTGATTCTTGAGTCTTGA
CCTCCGAATCTAGCTGGATTATTTCGACACACCAGACCACGTATCAGGT
TGCTCATCCCGAAATACTGCTTTGCAAACTGTTGTATCATCGCCTAGGAA
ATTAAGTTTCTTTTTTGGCTCTGTTACTGAATCAGTAGCTTTGCAACTTG
CTCATTATAAGCTGATCCATATTTTACATATCTTTTGAAGAATAATAGGT
ACTGACTTTACCTTTCTGATGAGAGCGATTTAAGAGATACCTCTGTAAAA
TCCATTTTTGTGAAGGGATCTGGGTTAGTTTTTAAAGGATTTGCTACAAC
AGTATCCCACAAACGATCTATTTCCCATTTNACTCATCCGCTCAAGATCT
ATCCACCTTTATATATGTTAATTGGGAGTCTTCCATGGTGCAATGAATCT
AGGATGCATTTAGAAGCCCAATCCATTACAAGTTTTCATCCAATTTCATG
TGACAAGTTGTTGGTTACTATGTAGGTACTTCCACAATTAAGAATTTCCA
GCAATGGATGTTGTTAATGCCATTCTTAAACCAGTTGCCGAGACACTTAT
GGAACCTGTTAAGAAACATCTAGGCTACATCATTTCCAGCACAAAACATG
TGAGGGATATGAGTAACAAAATGAGGGAGTTGAACGCTGCAAGACATGCT
GAAGAAGACCACTTGGACAGGAACATAAGAACTCGTCTTGAGATTTCAAA
TCAAGTTAGGAGTTGGTTAGAAGAAGTAGAAAAGATCGATGCAAAAGTAA
AAGCCCTTCCTAGTGATGTCACCGCTTGTTGCAGTCTCAAGATCAAACAT
GAAGTCGGAAGGGAAGCCTTGAAGCTAATTGTGGAGATTGAAAGTGCCAC
AAGACAACACTCTTTGATCACCTGGACTGATCATCCCATTCCTCTGGGAA

AAGTTGATTCCATGAAGGCATCGATGTCCACAGCATCAACCGATTACAAT
GACTTTCAGTCAAGAGAAAAAACTTTTACTCAAGCATTGAAAGCACTTGA
ACCAAACAACGCTTCCCACATGATAGCGTTATGTGGGATGGGTGGAGTGG
GGAAGACCACAATGATGCAAAGACTAAAAAAAGTTGCTAAACAAAATAGA
ATGTTCAGTTATATGGTTGAGGCAGTTATAGGGGAAAAGACGGACCCAAT
TGCTATTCAACAAGCTGTAGCGGATTACCTTCGTATAGAGTTAAAAGAAA
GCACTAAACCAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAGGCCAAC
TCTGGAGAAGGTAAGAATAAATTCCTTGTAATACTTGATGACGTCTGGCA
GTCTGTTGATCTAGAAGATATTGGTTTAAGTCCTTTTCCAAATCAAGGTG
TCGACTTCAAGGTCTTATTGACTTCACGAGACGAACATGTTTGCACAGTA
ATGGGAGTTGGATCTAATTCAATTCTTAATGTGGGACTTCTAATAGAAGC
AGAAGCACAAAGTTTGTTCCAACAATTTGTAGAAACTTCTGAGCCCGAGC
TCCATAAGATAGGAGAAGATATTGTAAGGAAGTGTTGCGGTCTACCTATT
GCCATCAAAACCATGGCATGTACTCTTAGAAATAAAAGAAAGGATGCTTG
GAAGGATGCACTTTCGCGTATAGAGCACTATGACCTTCGCAATGTTGCGC
CTAAAGTCTTTGAAACGAGCTACCACAATCTCCATGACAAAGAGACTAAA
TCAGTGTTTTTGATGTGTGGTTTGTTTCCGGAAGACTTCAATATTCCTAC
TGAGGAGTTGATGAGGTATGGATGGGGATTAAAGATATTTGATAGAGTCT
ATACATTTATAGAAGCAAGAAACAGGATCAACACCTGCATTGAGCGACTG
GTGCAGACAAATTTGTTAATTGAAAGTGATGATGTTGGGTGTGTCAAGAT
GCATGATCTGGTCCGTGCTTTTGTTTTAGGTATGTATTCTGAAGTAGAGC
ATGCTTCAGTTGTCAACCATGGTAATATACCTGGATGGACTGAAAATGAT
CCGACTGACTCTTGTAAAGCAATTTCATTAACATGCGAGAGTATGTCTGG
AAACATTCCAGGAGACTTCAAGTTTCCAAACCTAACGATTTTGAAACTTA
TGCATGGAGATAAGTCGCTAAGATTTCCACAAGACTTTTATGAAGGAATG
GAAAAGCTCCAGGTTATATCATACGATAAAATGAAGTATCCAATGCTTCC
CTTGTCTCCTCAATGCTCCACCAACCTTCGAGTGCTTCATCTCCATGAAT
GTTCATTAAAGATGTTTGATTGCTCTTGTATTGGAAATATGGCGAATGTG
GAAGTGTTGAGCTTTGCTAATTCTGGCATTGAAATGTTACCTTCCACTAT
CGGAAATTTAAAGAAGCTAAGGTTACTTGATTTAACAGATTGTCATGGTC
TTCATATAACACACGGTGTCTTTAACAATTTGGTCAAACTTGAAGAGTTG
TATATGGGATTTTCTGATCGACCTGATCAAACTCGTGGTAATATTAGCAT
GACAGATGTCAGCTACAATGAATTAGCAGAACGTTCAAAAGGCCTTTCTG
CATTAGAGTTCCAGTTCTTTGAAAACAATGCCCAACCAAATAATATGTCG
TTTGGGAAACTTAAACGATTCAAGATCTCAATGGGATGCACTTTATATGG
AGGATCAGATTACTTTAAGAAAACGTATGCTGTCCAAAACACATTGAAGT
TGGTTACTAACAAAGGTGAACTATTGGACTCTAGAATGAACGAGTTGTTT
GTTGAAACAGAAATGCTTTGTTTAAGTGTTGATGATATGAATGATCTTGG
TGATGTTTGTGTGAAGTCCTCACGTTCTCCTCAACCTTCTGTGTTCAAAA
TTCTAAGAGTCTTTGTCGTTTCCAAGTGTGTTGAGTTGAGATACCTTTTC
ACAATTGGTGTAGCCAAGGATTTGTCAAATCTTGAGCATCTTGAAGTTGA
TTCATGTAATAATATGGAACAACTCATATGTATTGAGAATGCTGGAAAAG
AGACAATTACATTCCTAAAGCTGAAGATTTTATCTTTGAGTGGGCTACCA

AAGCTTTCGGGTTTGTGCCAAAATGTCAACAAACTTGAGCTACCACAACT
CATAGAGTTGAAACTTAAGGGCATTCCAGGGTTCACATGCATTTATCCGC
AAAACAAGTTGGAAACATCTAGTTTGTTGAAGGAAGAGGTAGATATATGT
TTTATGTTAATACAAGTTAAAAAATCTTTTTAACTAAAAGTTTCAGTATA
TATATCTATATGTCTATAATTTGATTATATGATGTATTAGTGTTTGGATG
TGGCTATTAAGGGATGATTATTTTGCAGGTTGTGATTCCTAAGTTGGAGA
CACTTCAAATTGATGAGATGGAGAATTTAAAGGAAATATGGCATTATAAA
GTTAGTAATGGTGAGAGAGTTAAGTTGAGAAAGATTGAAGTGAGTAACTG
TGATAAGCTTGTGAATCTATTTCCACACAACCCCATGTCTCTGCTGCATC
ATCTTGAAGAGCTTGAAGTCAAGAAATGTGGTTCCATTGAATCGTTATTC
AACATCGACTTGGATTGTGTTGATGCCATAGGAGAAGAAGACAACATGAG
GAGCTTAAGAAACATTAAAGTGAAGAATTCATGGAAGTTAAGAGAAGTGT
GGTGTATAAAAGGTGAAAATAACTCTTGCCCCCTTGTTTCTGGCTTTCAA
GCTGTTGAAAGCATAAGCATTGAAAGTTGTAAGAGGTTTAGAAATGTATT
CACACCTACCACCACCAATTTTAATATGGGGGCACTTTTGGAGATATCAA
TAGATGACTGTGGAGAATACATGGAAAATGAAAAATCGGAAAAGAGTAGC
CAAGAGCAAGAGCAGGTATGGATTTCAATTTCACTTTCTTACTTACTTAA
GGATTAAGCTTCTGTTTTTTTGAATAAAAAAGGGACATCTTCTAATAATG
CACATCTTAAATTAAAAAGTATTTAATTGTTGCATAGCAGCGTATAACAT
CTTCTAATAATTTATCTGAAGGTGAAAGATCCAACTACTTCTAATTTGTT
AACAATTTCAATCATTTGCAAATGTTCCTTAAAAAATTAATTACCTGAAA
TCAAAACAATCTTCTTCAAATCCAAAATTATGAGACAGAATTGAGAAGGG
ATGTGAAATTATAAACCATTAACACAATTCCATGCTCACGTTACTAATTA
CATTTCTTGTTGGGATATATATGTACAGACTGATATTTTGTCAGAGGAAG
TGAAATTACAAGAAGTCACTGATACTATTTCTAATGTTGTATTCACATCG
TGTCTCATACACTCTTTTTATAACAACCTCCGTAAACTCAACTTGGAGAA
GTATGGAGGAGTTGAGGTTGTGTTTGAGATAGAGAGTTCAACAAGTAGAG
AATTGGTAACAACATACCATAAACAACAACAACAACAACAACCTATATTT
CCCAACCTTGAGGAATTATATCTATATTATATGGACAACATGAGTCATGT
ATGGAAGTGCAACAACTGGAATAAATTTTTACAACAATCAGAATCCCCAT
TCCACAACCTCACAACCATACACATGTCCGATTGCAAAAGCATTAAGTAC
TTGTTTTCACCTCTCATGGCAGAACTTCTTTCCAACCTAAAGAGAATCAA
TATTGACGAGTGTGATGGTATTGAAGAAATTGTTTCAAAAAGAGATGATG
TGGATGAAGAA

RG2O deduced polypeptide sequence (SEQ ID NO:119)

MDVVNAILKPVAETLMEPVKKHLGYIISSTKHVRDMSNKMRELNAARHAEEDHLD
RNIRTRLEISNQVRSWLEEVEKIDAKVKALPSDVTACCSLKIKHEVGREALKLIVEIE
SATRQHSLITWTDHPIPLGKVDSMKASMSTASTDYNDFQSREKTFTQALKALEPNN
ASHMIALCGMGGVGKTTMMQRLKKVAKQNRMFSYMVEAVIGEKTDPIAIQQAVA
DYLRIELKESTKPARADKLREWFKANSGEGKNKFLVILDDVWQSVDLEDIGLSPFP
NQGVDFKVLLTSRDEHVCTVMGVGSNSILNVGLLIEAEAQSLFQQFVETSEPELHKI

GEDIVRKCCGLPIAIKTMACTLRNKRKDAWKDALSRIEHYDLRNVAPKVFETSYHN
LHDKETKSVFLMCGLFPEDFNIPTEELMRYGWGLKIFDRVYTFIEARNRINTCIERL
VQTNLLIESDDVGCVKMHDLVRAFVLGMYSEVEHASVVNHGNIPGWTENDPTDSC
KAISLTCESMSGNIPGDFKFPNLTILKLMHGDKSLRFPQDFYEGMEKLQVISYDKMK
YPMLPLSPQCSTNLRVLHLHECSLKMFDCSCIGNMANVEVLSFANSGIEMLPSTIGN
LKKLRLLDLTDCHGLHITHGVFNNLVKLEELYMGFSDRPDQTRGNISMTDVSYNE
LAERSKGLSALEFQFFENNAQPNNMSFGKLKRFKISMGCTLYGGSDYFKKTYAVQ
NTLKLVTNKGELLDSRMNELFVETEMLCLSVDDMNDLGDVCVKSSRSPQPSVFKIL
RVFVVSKCVELRYLFTIGVAKDLSNLEHLEVDSCNNMEQLICIENAGKETITFLKLKI
LSLSGLPKLSGLCQNVNKLELPQLIELKLKGIPGFTCIYPQNKLETSSLLKEEVVIPKL
ETLQIDEMENLKEIWHYKVSNGERVKLRKIEVSNCDKLVNLFPHNPMSLLHHLEEL
EVKKCGSIESLFNIDLDCVDAIGEEDNMRSLRNIKVKNSWKLREVWCIKGENNSCPL
VSGFQAVESISIESCKRFRNVFTPTTTNFNMGALLEISIDDCGEYMENEKSEKSSQEQ
EQTDILSEEVKLQEVTDTISNVVFTSCLIHSFYNNLRKLNLEKYGGVEVVFEIESSTS
RELVTTYHKQQQQQQPIFPNLEELYLYYMDNMSHVWKCNNWNKFLQQSESPFHN
LTTIHMSDCKSIKYLFSPLMAELLSNLKRINIDECDGI

RG2P polynucleotide sequence (SEQ ID NO:120)

CCCATTGCTATTCAGGAAGCAGTAGCAGATTACCTCNGTATAGAGCTCAA
AGAAAAAACTAAATCNGCAAGAGCTGATATGCTTCGTAAAATGTTAGTTG
CCAAGTCCGATGGTGGTAAAAATAAGTTCCTAGTAATACTTGACGATGTA
TGGCAGTTTGTTGATTTAGAAGATATCGGTTTAAGTCCTTTGCCAAATCA
AGGTGTTAACTTCAAGGTCTTGCTAACATCACGGGATGTAGATGTTTGCA
CTATGATGGGAGTCGAAGCCAATTCAATTCTCAACATGAAAATCTTACTA
GATGAAGAAGCACAAAGTTTGTTCATGGAGTTTGTACAAATTTCGAGTGA
TGTTGATCCCAAGCTTCATAAGATAGGAGAAGATATTGTAAGAAAGTGTT
GTGGTTTGCCTATTGCCATCAAAACCATGGCCCTTACTCTTAGAAATAAA
AGCAAGGATGCATGGAGTGATGCACTTTCTCGTTTAGAGCATCATGACCT
TCACAATTTTGTGAATGAAGTTTTTGGAATTAGCTACGACTATCTTCAAG
ACCAGGAGACTAAATATATCTTTTTGCTTTGTGGATTGTTTCCCGAAGAC
TACAATATTCCTCCTGAGGAGTTAATGAGGTATGGATGGGGCTTAAATTT
ATTTAAAAAAGTGTATACTATAAGAGAAGCAAGAGCCAGACTCAACACCT
GCATTGAGCGGCTTATCCATACCAATTTGTTGATGGAAGGAGATGTTGTT
GGGTGTGTAAAGATGCATGATCTAGCACTTGCTTTTGTTATGGATATGTT
TTCTAAAGTGCAGGATGCTTCAATTGTCAACCATGGTAGCATGTCAGGGT
GGCCTGAAAATGATGTGAGTGGCTCTTGCCAAAGAATTTCATTAACATGC
AAGGGTATGTCTGGGTTTCCTATAGACCTCAACTTTCCAAACCTCACAAT
TTTAAAACTTATGCATGGAGATAAGTTTCTCAAGTTTCCTCCAGACTTTT
ATGAACAAATGGAAAAGCTTCAAGTTGTATCGTTTCATGAAATGAAATAT
CCGTTTCTTCCCTCGTCTCCTCAATATTGCTCCACCAACCTTCGAGTTCT
TCATCTCCATCAATGCTCATTGATGTTTGATTGCTCTTGTATTGGAAATC
TGTTTAATCTGGAAGTGTTGAGCTTTGCTAATTCTGGCATTGAATGGTTA

CCTTCCAGAATTGGAAATTTGAAGAAGCTAAGGCTACTAGATTTGACAGA
TTGTTTTGGTCTTCGTATAGATAAGGGTGTCTTAAAAAATTTGGTCAAAC
TTGAAGAGGTTTATATGAGAGTTGCTGTTCGAAGCAAAAAAGCCGGAAAT
AGAAAAGCCATTAGCTTCACAGATGATAACTGCAATGAGATGGCAGAGCG
TTC

RG2P deduced polypeptide sequence (SEQ ID NO:121)

PIAIQEAVADYL?IELKEKTKSARADMLRKMLVAKSDGGKNKFLVILDDVWQFVDL
EDIGLSPLPNQGVNFKVLLTSRDVDVCTMMGVEANSILNMKILLDEEAQSLFMEFV
QISSDVDPKLHKIGEDIVRKCCGLPIAIKTMALTLRNKSKDAWSDALSRLEHHDLHN
FVNEVFGISYDYLQDQETKYIFLLCGLFPEDYNIPPEELMRYGWGLNLFKKVYTIRE
ARARLNTCIERLIHTNLLMEGDVVGCVKMHDLALAFVMDMFSKVQDASIVNHGS
MSGWPENDVSGSCQRISLTCKGMSGFPIDLNFPNLTILKLMHGDKFLKFPPDFYEQ
MEKLQVVSFHEMKYPFLPSSPQYCSTNLRVLHLHQCSLMFDCSCIGNLFNLEVLSF
ANSGIEWLPSRIGNLKKLRLLDLTDCFGLRIDKGVLKNLVKLEEVYMRVAVRSKKA
GNRKAISFTDDNCNEMAERS

RG2Q polynucleotide sequence (SEQ ID NO:122)

TGGGGAAGACACAGTGATAGAAAARAAAAAGAATGTTGTGGAAAAGAGGA
AAATGTTTGATTATGCTGTTGTGGCGGTTATAGGGGAAAAGACGGACCCT
ATTGCTCTTCAGAAAACTGTTGCGGATTACTTGCATATTGAGCTAAATGA
AAGCACTAAACTAGCAAGAGCAGATAAACTTTGCAAATGGTTCAAGGACA
ACTCGGATGGAGGTAAGAAAAAGTTCCTCGTAATACTCGACGATGTTTGG
CAATCTGTTGATTTGGAAGATATTGGTTTAAGTACTCCTTTTCCAAATCA
AGGTGTCAACTTCAAGGTTTTGTTGACATCACGAAAGAGAGAAATTTGCA
CAATGATGGGAGTTGAAGCTGATTTAATTCTCAATGTCAAAGTCTTAGAA
GAAGAAGAAGCACAAAAGTTGTTCCTCCAGTTTGTAGAAATTGGTGACCA
ATACCACGAGCTTCATCAGATAGGGGTACATATAGTAAAGAAGTGTTATG
GTTTACCCATTGCCATTAAAACCATGGCTCTTACTTTAAGAAATAAAAGA
AAGGATTCATGGAAGGACGCACTCTCTCGTTTAGAGGACCATGACACTGA
AAATGTTGCAAATGCAGTTTTCGAGATGAACTACCGCAATCTACAAGATG
AGGAGACCAAAGCCATTTTTTTGCTTTGCGGTTTGTTCCCCGAAGACTTT
GATATTCCTACTGAGGAGTTGGTGAGGTATGGATGGGGCTTAAATCTATT
TAAAAAAGTGTATACCATAAGAAAGGCAAGAACGAGATCGCATACATGTA
TTGAGCGACTCTTGGATTCAAATTTGTTGATTGAAAGTAACGATATTCGG
TGCGTCAAGATACACGATCTGGTGCGCGCTTTTGTTTTGGATATGTATTG
TAAAGTTGAGCATGCTTCAATTGTCAACCATGGTAATATGCGGACCGAAT
ATAATATGGCTGACTCTTGCAAAACAATTTCATTAACATACAAGAGTATG
TCTGGGTTTGAGTTTCCAGGAGACCTCAAGTTTCCAAACCTAACAGTTTT
GAAACTTATGCANGGAGATAAGTCTCTAAGGTTTCCTCAAGACTTTTATC
AATCAATGGAAAAACTTCGGGTTATATCATATGATAAAATGAAGTATCCA
TTGCTTCCCTCATCACCTCAATGCTCCACTAACATCCGAGTGCTTCGTCT

CCATGAATGTTCATTAAGGATGTTTGATTGCTCTTGTATTGGAAAGCTAT
TGAATTTGGAAGTCCTCAGCTTTTTAATTCTAACATTGAATGGTTACCT
TCCACAATCAGAAATTTAAAAAGCTAAGGCTACTAGATTTGAGATATTG
TGATCGTCTTCGTATAGAACAAGGTGTCTTGAAAAATTTGGTCAAACTTG
AAGAACTTTATACTGGATATACATCAGCGTTTACAGA

RG2Q deduced polypeptide sequence (SEQ ID NO:123)

GEDTVIEKKKNVVEKRKMFDYAVVAVIGEKTDPIALQKTVADYLHIELNESTKLAR
ADKLCKWFKDNSDGGKKKFLVILDDVWQSVDLEDIGLSTPFPNQGVNFKVLLTSR
KREICTMMGVEADLILNVKVLEEEEAQKLFLQFVEIGDQYHELHQIGVHIVKKCYG
LPIAIKTMALTLRNKRKDSWKDALSRLEDHDTENVANAVFEMNYRNLQDEETKAI
FLLCGLFPEDFDIPTEELVRYGWGLNLFKKVYTIRKARTRSHTCIERLLDSNLLIESN
DIRCVKIHDLVRAFVLDMYCKVEHASIVNHGNMRTEYNMADSCKTISLTYKSMSG
FEFPGDLKFPNLTVLKLM?GDKSLRFPQDFYQSMEKLRVISYDKMKYPLLPSSPQCS
TNIRVLRLHECSLRMFDCSCIGKLLNLEVLSFFNSNIEWLPSTIRNLKKLRLLDLRYC
DRLRIEQGVLKNLVKLEELYTGYTSAFTE

RG2S polynucleotide sequence (SEQ ID NO:124)

ATTTGGGGTTTTACATTTAATTTTTTGTGCATGAATGTGAAAATAGACTG
CTTATTGATTCTTTGTGTTTCATTGAGTTGATTTTCATTATTACTACCTT
ACAAATTGCTCAGTGATAGATTTCCATTAATTTGCTAATTCGGTTGCTTC
TAAATATGTAGGAGCTACTAAAAGCAAAAATATCGAGCAATGTCGGACCC
AACGGGGATTGCTGGTGCCATTATTAACCCAATTGCTCAGAGGGCCTTGG
TTCCCGTTACAGACCATGTAGGCTACATGATTTCCTGCAGAAAATATGTG
AGGGTCATGCAGACGAAAATGACAGAGTTGAATACCTCAAGAATCAGTGT
AGAGGAACACATTAGCCGGAACACAAGAAATCATCTTCAGATTCCATCTC
AAATTAAGGATTGGTTGGACCAAGTAGAAGGGATCAGAGCAAATGTGGAA
AACTTTCCGATTGATGTCATCACTTGTTGTAGTCTCAGGATCAGGCACAA
GCTTGGACAGAAAGCCTTCAAGATAACTGAGCAGATTGAAAGTCTAACAA
GACAGCTCTCCCTGATCAGTTGGACTGATGATCCAGTTCCTCTAGGAAGA
GTTGGTTCCATGAATGCATCCACCTCTGCATCATCAAGTGATGATTTCCC
ATCAAGAGAGAAAACTTTTACACAAGCACTAAAAGCACTCGAACCCAACC
AACAATTCCACATGGTAGCCTTGTGTGGGATGGGTGGAGTAGGGAAGACT
AGAATGATGCAAAGGCTGAAGAAGGCCGCTGAAGAAAAGAAATTGTTTAA
TTATATTGTTAGGGCAGTTATAGGGGAAAAGACGGACCCCTTTGCCATTC
AAGAAGCTATAGCAGATTACCTCGGTATACAACTCAATGAAAAAACTAAG
CCAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAAAAGAATTCAGATGG
AGGTAAGACTAAGTTCCTCATAGTACTTGACGATGTTTGGCAATTAGTTG
ATCTTGAAGATATTGGGTTAAGTCCTTTTCCAAATCAAGGTGTCGACTTC
AAGGTCTTGTTGACATCACGAGACTCACAAGTTTGCACTATGATGGGGGT
TGAAGCTAATTCAATTATTAACGTGGGCCTTCTAACTGAAGCAGAAGCTC
AAAGTCTGTTCCAGCAATTTGTAGAAACTTCTGAGCCCGAGCTCCAGAAG

ATAGGAGAGGATATCGTAAGGAAGTGTTGCGGTCTACCTATTGCCATAAA AACCATGGCATGTACTCTTAGAAATAAAAGAAAGGATGCATGGAAGGATG CACTTTCGCGCATAGAGCACTATGACATTCACAATGTTGCGCCCAAAGTC TTTGAAACGAGCTACCACAATCTCCAAGAAGAGGAGACTAAATCCACTTT TTTAATGTGTGGTTTGTTTCCCGAAGACTTCGATATTCCTACTGAGGAGT TGATGAGGTATGGATGGGGCTTGAAGCTATTTGATAGAGTTTATACGATT AGAGAAGCAAGAACCAGGCTCAACACCTGCATTGAGCGACTGGTGCAGAC AAATTTGTTAATTGAAAGTGATGATGTTGGGTGTGTCAAGATGCATGATC TGGTCCGTGCTTTTGTTTTGGGTATGTTTTCTGAAGTCGAGCATGCTTCT ATTGTCAACCATGGTAATATGCCCGAGTGGACTGAAAATGATATAACTGA CTCTTGCAAAAGAATTTCATTAACATGCAAGAGTATGTCTAAGTTTCCAG GAGATTTCAAGTTTCCAAACCTAATGATTTTGAAACTTATGCATGGAGAT AAGTCGCTAAGGTTTCCTCAAGACTTTTATGAAGGAATGGAAAAGCTCCA TGTTATATCATACGATAAAATGAAGTACCCATTGCTTCCTTTGGCACCTC GATGCTCCACCAACATTCGGGTGCTTCATCTCACTAAATGTTCATTAAAG ATGTTTGATTGCTCTTGTATTGGAAATCTATCGAATCTGGAAGTGCTGAG CTTTGCTAATTCTCGCATTGAATGGTTACCTTCCACAGTCAGAAATTTAA AGAAGCTAAGGTTACTTGATCTGAGATTTTGTGATGGTCTCCGTATAGAA CAGGGTGTCTTGAAAAGTTTAGTCAAACTTGAAGAATTTTATATTGGAAA TGCATCTGGGTTTATAGATGATAACTGCAATGAGATGGCAGAGCGTTCTG ACAACCTTTCTGCATTAGAATTCGCGTTCTTTAATAACAAGGCTGAAGTG AAAAATATGTCATTTGAGAATCTTGAACGATTCAAGATCTCAGTGGGACG CTCTTTTGATGGAAATATCAATATGAGTAGCCACTCATACGAAAACATGT TGCAATTGGTGACCAACAAAGGTGATGTATTAGACTCTAAACTTAATGGG TTATTTTTGAAAACAAAGGTGCTTTTTTTAAGTGTGCATGGCATGAATGA TCTTGAAGATGTTGAGGTGAAGTCGACACATCCTACTCAGTCCTCTTCAT TCTGCAATTTAAAAGTTCTTATTATTTCAAAGTGTGTAGAGTTGAGATAC CTTTTCAAACTCAATCTTGCAAACACTTTGTCAAGACTTGAGCATCTAGA AGTTTGTGAATGCGAGAATATGGAAGAACTCATACATACTGGAATTTGTG GAGAAGAGACAATTACTTTCCCTAAGCTGAAGTTTTTATCTTTGAGTCAA CTACCGAAGTTATCAAGTTTGTGCCATAATGTCAACATAATTGGGCTACC ACATCTCGTAGACTTGATACTTAAGGGCATTCCAGGTTTCACAGTCATTT ATCCGCAGAACAAGTTGCGAACATCTAGTTTGTTGAAGGAAGAGGTAGAT ATATGTTCTTTATGTTAATACAATTTAAATAATATTTTCAACCAAATTTT CATAATATATCTGTAATTTGATTGTATGATGTGTTATTGTTTATATGTGG CTATTAAGGGATGATTATTTTGCAGGTTGTGATTCCTAAGTTGGAGACAC TTCAAATTGATGACATGGAGAACTTAGAAGAAATATGGCCTTGTGAACTT AGTGGAGGTGAGAAAGTTAAGTTGAGAGAGATTAAAGTGAGTAGCTGTGA TAAGCTTGTGAATCTATTTCCGCGCAATCCCATGTCTCTGTTGCATCATC TTGAAGAGCTTAAAGTCAAGAATTGCGGTTCCATTGAATCGTTATTCAAC ATTGACTTGGATTGTGTCGGTGCAATTGGAGAAGAAGACAACAAGAGCCT CTTAAGAAGCATCAACATGGAGAATTTAGGGAAGCTAAGAGAGGTGTGGA GGATAAAAGGTGCAGATAACTCTCATCTCATCAACGGTTTTCAAGCTGTT

GAAAGCATAAAGATTGAAAAATGTAAGAGGTTTAGCAATATATTCACACC TATCACCGCCAATTTTTATCTGGTGGCACTTTTGGAGATTCAGATAGAAG GTTGCGGAGGAAATCACGAATCAGAAGAGCAGGTAACGCTTTCAATTTAA CTTTCTTAAGTAATTAAGGACTAACCTCCTGTTTTTGAATAATAAAGAG GTGGGATGACTAAACTTGGGCATCACAATTGCAACAAAATGTTACAAACC ATGAAACGTTCAAACCATTTCTTGAATTAAGGTTTCAATACAAGTCATTT AAAAATATGGCTTAAATTTTTTATATTTATGTATCAACATGATTTTTCA TTAGAGATCATTATTATAATAGTAAGTTTAAAGCAATTTAAATTAGAACT AATTCTAACTTTAGCTAATAAATCGTTATAAATGTAAATAATTACTTTTT AGTGAAATAAGCAACGGATTTAATAAGTTAACAACTTAAATGTCATTTCC TAACAAAAAAAACTATTTGGTTCAGAAGAACCGTAATTCAAGATAACTAA AATAAAAATATTTGACATTCACTAAGAGCATTTTTTTTCTAAATATGAT TGCAAATGAATAAAACTTAAATTTATACAGAAAAGATTTTTATATATGTT ATACAAAATTTACAAATTGAAACTGGATATGTTAATTAACGGTTTATAAT TCTGGTATCACAAAGGGATATATAATAAAATATTATTTCTGTAGTCATT TATAATTGTACTAGTTTATAACCCGTGGGAACCATGAGTTCTAAAATTAG TTAAACTTTCATAATAAAAATTTATAATTATTATTTATTTAAATAAATT ATTAATTAAGAGATGTATCAAAAATTTAAAGTTATTATAACTTCAAATTT AACATATAATTAGAAAATATATGATCATAACTTTCCGCAACTCTTCTTTT GTATTAAAATGCCCAGAGAAGCTCTTAGTAYATTTCTAAATCAAAGTCA CAAAACTAATGAAGCATATAATTTTGTGAAAATCAATTAGCATTAGGTTT TAAGAGTCACCAAATTCAAAGAGTAATCCAATGCTTTCATTACCACTATG GAGAAAATATTTTCTTAGTTTAAATGAAATGAAAACAAACATTCAAACTA ATTGTTGCTTACTAAACCAAAGACCCATTACTTAGCCAAGAGTTTAACCA AAAAAAATTACATTCATGTATCATTATTCATGACTAGATATATATGAACA TGAAGGGAGTTTTTATAGAAAATATAATCATAGATATTCAACATAACTTC ATGGAATTCCTCAAAATAACCAAGTTATTCAAGAAATTACATCCAAGTCA ACCAAAGAGAAGTTTAGCCTAGCATGGCTAAACTCAAGAAAATAAAATAA GGATTAGAAGTACCAAACATGTAGTAAGAATCACAGTAAAAGATGATGTT GTTCTTGATGTTCTTCTAAGTTCTTCAAGTCTCCAGTTGCTCCTAATAAT GCAAAGGAGAGCCATTAAATTCGTATGTATTGATCCCTTCAAAAGCTGCA CCAACCTCCCTTAAATAACACTCAAAGCAAAAATGACAAAATTGCCCCTG AAGGACCCTATGCGGGTGCCTTGCGCGGGTGGAGCTGAATACGAAAGGTC TTTGGTCTTTGTGAGGGTGATGCTGTGCGGGTTAGCTTGTCGCATGCTTC CGCGCGGTTCGCGCACATGTGCACAAGTGATGCATGGTGTGTACGTTCTT GAGTTTTGAGCCTCCGATGCTTAGTCCATTTGGCCCAATTCGAGTCCAAT CAGCTTATGACCCATTTTTCTTCAAGTTATCTTCAAGTTATCTTCAAGTT AAGCCCAAATTGCCTTCTCCAAATCATCCATAACTTCACAAAATCGCCCG TTCATCTTAATCCCGAATGCACAATTATTCTCCTGTCTTCCTTTTAAGCA AGATACCACCTTCTTCATGCTTCATCCATCAATAGTACACTTCATGTATC ATCTCTACTAGTTATTTAGTCCACAATCCTTATTGTCCTCCAAATTTAAT TATCTCATTTAGTTCCCGTTCCACTAGTTTCCTTAAAATTTGCAATTAAG CTCACACAAATATTAAGTACCTGAAATGGTCATAAAATAACAAAAAGGAA

AATATGCATGAAGATTAACTAAATGATGAACGAAATATGCTAAAATAGAC
TATAAAATGAAGTAAATAAAATGAAATTATCGCACTCCGACCACCCTTAT
AGGCTTGTAGTCCATCCACCCTTCATTCCTTGTACCAATATGGGATGGAA
ACATCATTAATTAAGCCAAAAAACTAACATATAAGGGGTGAGTGACAAAG
GTAAGTACTAAAGATGAAAATAATCCATTTTTYTTGTATATACACAACAC
ACACATAGGGGCAGACGTAGGATTTCATAGTACAGATTGTTGGTGGCACA
TAAGTGTTGCTGGTGACACTTTTTTTTTCTTTTACGTAGTGGCACAACAG
TAGAAAAAACGARAAATTCGAAATTTTTTACAATGTGTSTAAAAAAAAYA
GTGGTTGTTGGTGCCACTATGGACACCAAAGTTGAACTGCCCCTGCGCGC
RCACACACACACACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
ARAGWAWGRRRGAKAKARMCSMSYTTGGGATGTGATACTTCTTTTAGGAA
AATGGAGTTATATCTTTGATATTGTATTTTTTTAATGTAATTTATATATT
TAATCATTTTAGTTTATAAGTTTTATTTATTTTGATATGAAAAAAAAAGT
CTTTTATACATTGGATTTAACATAAAAATCCAACAATATTAATCAAAAAG
ACCAMACATGTGGACAMWTATGTATATAAWTAATTCACAATAGTCTTTAG
GAATAGNATTATATATATAATTAATTCTCAATGGTCTTAGGAATAGTAAG
TTCTTATATTTCAAACTTTNGCCACAATTCTTTGKTTACTTWGACACTTY
CCTCTCTCTAATTATATATATATATATATATATATATATATATATACACA
CACACACACACACACTAGATGTGTGCCCGCGCAAAGCAGTGACGTNNNGG
AGAANACTTTCTTAAGCATAAATAATTATTATATTTTTTATTGGGTATTA
TATAATAAAAAATTACAACTTTTAAATAAAATATTTATGTTTATACTTTA
TATTTATATTGCTTGTATACTATTAATATAATAAATTAATATTTATGTCT
AATTTATGAAATGTAAATTAATTTAAATACATGAATTTAATATTTTTAAA
ATTTTCAGTTTGCTTCAAATTGAGTTTCTTAATTATTTTTTTTAATTCAN
GTATTCAAACTTTTGGTAAGTATTAAAGAATTATTTATGCACAATTGATT
TATACAAAAAACTTTGTAACTTATACATCTTAAAATTCAAGATATAACTA
ACATGTTTTACAATATATATATATATATANATATATATATATATATATATAT
ATATATATATATATATAGTAAAGCGCANAGGTCATAGGNANAGANTATTT
TCTATTATTCTACGTTTTGCCACAAAAGTTTGAACACTTTGCCACTTTTT
GTCCCTCCTTAACCTTTTCAATGTTTTGCGACAAAAGTTCCAAAACTTTG
CCACTTTGATCATTCCTCAACTTTTCACCGCATTAGTTTGTGGAGTTGGC
AGTTTTGGTCCCCCTAACTTCGATATTTTCTCCTGCTAGCCAAAAAGGGT
TCCAGAGTTTCACANTTTTGGTCCCTGACAATAACCAAATGTGAGATGTC
AAATTTTTGCCACATTAGTTTGTGGAGTTGTCCCTTTTGGTCCCCCCACA
TTCGATATTCTACTATACGACCTTATTTTTCTCAAATAACAACACGTATA
TTTAATTACCAATGATAGAAATAGATATCAAATAAAGTATTTGTAACACC
GTGTAAGAACGGTGCTACTATAGGTAAAAATAAACATTTCAAAGTACGAT
GTCCTAATTGGAAAAAGAGTTTTAAAAAAATAACAACTAGGGGCGAGTTT
TTTTTACAAGTTTGTATCAAATCATATCAAAATTTAAGGTGGAACGGTGA
CCACATTAACCAGAAATGTAATTTATTCTTTGATTTTGATAATTTTTAAT
ATTTTGTTGTGATCTATGTATTTAAAAGTAAACAACAAAGAACATAATCC
AAAACCCTAAATTGCAAGTCTCGCCCAATTTCTCTATCACTAGTCGTCAC
TTACGATGGCGTTACGTCGCTCTCTCACTTCTTACAACCCTTTGTTGCTA

CTCATTACAATAACGAAAAGTTGAATATCCATATATTTATTTGGATGTGG
AATTGAACAAATCTCGTCAAATTTTTGATTTTGTTGATGGATTTGAGTAG
AAGTTTGGGCAGAACGGGAATGATGGTCTGCAAGTGGTTATAAACTTGAT
TCTGAGTTATTACTATATATGTAGCCTCTTTACAACGACCAAGGTTTCTT
CCAGGTACCATTTGATCTTTTTAGAACCCAGTTGTCTGAAACACCCTGAT
TTGGATCAAATATCACCAACAACTCTTAAAAACTTGATTAATCAATTGTT
TTCTTCATCTTGATAACAAGTGGAATGATTTCTACTTAGATTAACTTGA
AAAAAAAGGTCCATGTGCGTCTGGTGGATCTGGTAAATGAAGATGGAAGG
GAGAGCTGACTTTAAAGACACAAACACGTCACCATATCTTTTATTTTATT
TTAAATTTGCTTTTTTCCTATTTCTTTCTTTCTTGATCTCCAGATGGTAT
GTGGTGTGGATAATTTACACATAGAGATTGGGAACGACTGTGTTTTAGAG
AGGACGTGGCTTGGGGTTGAGGATGGTTTATGGCTGGCCGAGTTTCATTT
ATATAAACAAACAAATATATAAAACAAGGGGTAAAATGGCCATCTTATAT
GTATTTAACCGTCCTTTTTTATTTTTTTTTTATTTTTAAATTTAAGAAGG
GGTATACCAGTGTCAGCCTCTTATTCCCAACCAGGCAACCAGTCAAATAG
GGACTTAGGTTGTTTGGAAACAGTTCCGTGAGACCGTGACTTGGATGGTA
GATAAATTTAGTAAACTTAACCCTTCAATTAACCTACCTTTTTCTTATTA
ACTCAATTTCAACCTAAATTCTGATTCTTGTTTGAAAATAAGTTGCATCT
TTATGTTTGTATTATCCTGTTGCATAGGATCCTTAGCATCTTTTAATAAT
TTATTTGAAGGTGAAAGATCCAACTATTTTTTAGCTGTTGGCATTTTCCA
TCATTTGCAACTGTTTCTTGAAAAAAAAATACCTAAAATCAAAATAACCA
TTTTCAAATCCAAAATTATAAGAGAGAATTGTTAATGGACGTGGAATCGT
AAATCATTAACACAGTTCAGTACACAAGTTGCTAATTACATTTCTTGCTG
TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGTCACTGAT
ACTAATATTTCTAATGATGTTGTATTATTCCCATCCTGTCTCATGCACTC
TTTTCATAACCTCCATAAACTTAAATTGGAGAGAGTTAAAGGAGTGGAGG
TGGTGTTTGAGATAGAGAGTGAGAGTCCAACAAGTAGAGAATTGGTAACA
ACTCACCATAACCAACAACATCCTATTATACTTCCCAACCTCCAGGAATT
GGATCTAAGTTTTATGGACAACATGAGTCATGTGTGGAAGTGCAGCAACT
GGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAATCCCCATTCCAC
AACCTCACAACCATACACATGTTCAGCTGCAGAAGCATTAAGTACTTGTT
TTCGCCTCTCATGGCAGAACTTCTTTCCAACCTAAAGGATATCTGGATAA
GTGGGTGTAATGGTATTAAAGAAGTTGTTTCAAAGAGAGATGATGAGGAT
GAAGAAATGACTACATTTACATCTACCCACACAACCACCATCTTGTTCCC
TCATCTTGATTCTCTCACTCTAAGACTACTGGAGAATCTGAAGTGTATTG
GTGGAGGTGGTGCCAAGGATGAGGGGAGCAATGAAATATCTTTCAATAAT
ACCACTGCAACTACTGCTGTTCTTGATCAATTTGAGGTATGCTTTGTACA
TATTCAATTATTTATTTAATTTCCTTTTTTCTTTGCAATATTCTATAAAT
AATACATTTTATACCCACTATACTAAGATAATAATTACCTAGAGGGATGG
ATGCTATGACACAGCTGCTACACTTCAGAAACTCTAGTAAGGGCAGTTAT
GGAAGTTCAATAAAATGATAATGGCATCTTTTGATGGGTAATATAGGCAA
TTTAAGTTTTATTTCTGTTAAAGCAGTATTTAGCAAGTACTGGCCAGTAG
GAGAGGAGAATATCACCTTTTGTGAAAATCTGGTCATTGTACCCAAGAAT

TTAGTTAAATGTAACATTTTAGATATCAGGGGACATCAGGTGACAGATAT
TGTAGAATAGAACAATATATAATATTACCCAAAACTATTTTTCTAAGGT
TATTCTGTTAAATATGTGCTTTCTTGATTTCATTGAATTTGCATTCCTAT
ATTTTAGGTGGTAAAGTGATTGTCTCTTCAATAAATCCCGAAATTAATTA
AAAAAAAAAAAAACAAAAGTAAATTTTTGATATGGAGAGCACTGGTATCA
TTTAGTATATAAAAAAACTAGATTTTGAATTAAGTTTCTTATATAAAAGC
TGTGTATATAGTTTAATTAGTTTTACATCATTTTTCCATGTGGTGTTGCA
GTTGTCTGAAGCAGGTGGTGTTTCTTGGAGTTTATGCCAATACGCTAGAG
AGATAGAGATATCTAAGTGTAATGTATTGTCAAGTGTGATTCCATGTTAT
GCAGCAGGACAAATGCAAAAGCTTCAAGTGCTGAGAGTAACGGGTTGTGA
TGGCATGAAGGAGGTATTTGAAACTCAATTAGGGACGAGCAGCAACAAAA
ACAGAAAGGGTGGTGGTGATGAAGGAAATGGTGGAATTCCAAGAGTAAAT
AACAATGTTATTATGCTTCCCAATCTAAAGACATTGAAAATCTACATGTG
CGGGGGTTTGGAACATATATTCACATTCTCTGCACTTGAAAGCCTGACAC
AGCTCCAAGAGTTAAAGATAGTGGGTTGCTACGGAATGAAAGTGATTGTG
AAGAAGGAAGAAGATGAATATGGAGAGCAGCAAACAACAACAACAACAAC
AACGAAGGGGGCATCTTCTTCTTCTTCTTCTTCTTCTAAGAAGGTTG
TGGTCTTTCCCCGTCTAAAGTCCATTGAACTATTCAATCTACCAGAGCTG
GTAGGATTCTTCTTGGGGATGAATGAGTTCCGGTTGCCTTCATTGGAAGA
AGTTACCATCAAGTATTGCTCAAAAATGATGGTGTTTGCAGCTGGTGGGT
CCACAGCTCCCCAACTCAAGTATATACACACAAGATTAGGCAAACATACT
CTTGATCAAGAATCTGGCCTTAACTTTCATCAGGTATATATATATTCCTT
TAATTGGCATGATCTAATTAAGAAAGATATCATTCCTGCCAAGTAAATTT
ACTTCAAACACATTCACACTGGTTTCAGTCTAAGTTTATGTTGTTCTAGG
AAGGCCAAAATGGGAAAGCAAGATAGGGAAAAATAGTGTATTTCAGTGGA
AAGGGTATTTTAGGTATTTTCTGTCAAAAGTTGTTATTGCAGGCTTTTTA
GTACCTGGAATCGTGTGTGGGAGGAGCGTTATTATTCTGATTTGCTTGTT
TCTTTATCATTTTTTCTTAGCCTCTCGAACAGCTAGAAACCCTTTTAATC
TTTTGATTTTAAATGACAAAATTTTTCCCTGTTACTCTATTTGATTGTTG
TTCTTCATGGTTCTAAGTGAGTTATTGGCTCATCTGTTACTTCTTTTGAT
TGTTATTTTCATATCATGTTGTCCTTTGAATCAAGCTTTTCCATTTTCAA
CCAGGGCAAAAGGTCAAAAGTAACCTACTTTATGAGATCAAAAACAGCAA
CCCATCGGATAACTTTTAGTTGGAGTTAATAGTTACAATTACCATTGTGA
TTAATAATTATAATATCTTGTATTAATTCATTAAAATTGGTACAGCACAT
ATATGACATTTTAAAGGTTTGTTTTTGTTWGACATATATATGCCTCTGGC
GTTTTCTTTATTGGACATGCAGACCTCATTCCAAAGTTTATACGGTGACA
CCTCGGGCCCTGCTACTTCAGAAGGGACAACTTGGTCTTTTCATAACTTG
ATCGAATTAGATATGGAATTAAATTATGATGTTAAAAAGATTATTCCATC
CAGTGAGTTGCTGCAACTGCAAAAGCTGGAAAAGATTCATGTGAGTAGTT
GTTATTGGGTAGAGGAGGTATTTGAAACTGCATTGGAAGCAGCAGGGAGA
AATGGAAATAGTGGAATTGGTTTTGATGAATCGTCACAAACTACTACTAC
TACTACTCTTTTCAATCTTCGAAACCTCAGAGAAATGAAGTTGCATTTTC
TACGTGGTCTGAGGTATATATGGAAGAGCAATCAGTGGACAGCATTTGAG

TTTCCAAACCTAACAAGAGTTCATATAAGTAGGTGTAGAAGGTTAGAACA
TGTATTTACTAGTTCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTAG
ATATTAGTTGGTGCAACCATATGGAGGAGGTGATTGTTAAGGATGCAGAT
GTTTCTGTTGAAGAAGACAAAGAGAGAGAATCTGATGGCAAGACGAATAA
GGAGATACTTGTGTTACCTCGTCTAAAATCCTTGAAATTAAAATGCCTTC
CATGTCTTAAGGGGTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATTA
TTGGATACTTTAGAAATCTACAAATGCCCAGCAATAACGACCTTCACCAA
GGGAAATTCTGCTACTCCACAGCTAAAAGAAATAGAAACAAGATTTGGCT
CGTTTTATGCAGGGGAAGACATCAACTCCTCTATTATAAAAAGATCAAAC
AACAGGTAAATCAGATCTTTGTTGCTTTAATAATTCTTAAACTACATTTG
AAAAGCTTCATGCAAGTTTTTTTTGTTATATTGTCAAAAACCGCAACCTA
CATTTTCAGCTTTATATTTATGTACTTTATGCAGGAGTTCAAACAAAACT
CTGATTAATGTGAAGTGAATATTAAAGGTAAATTATATTTTCATGTTCCT
AGTTGCCTATTAATTAATGGCCTTTTAGTTCRTGATTTTTGGATGTAGTY
WTCATGATGATGTGAATCTTCTAATACCCCATTCATTGTTTGGTTGAATG
TTGACTCTATGTCAGGATGAATATTCAAGGGAAGAATTGTTCATCATATG
AAGGACATTAAAGAACATGGATGCTATGAAGATGTTGGAARAC

RG2S deduced polypeptide sequence (SEQ ID NO:125)

MSDPTGIAGAIINPIAQRALVPVTDHVGYMISCRKYVRVMQTKMTELNTSRISVEEH
ISRNTRNHLQIPSQIKDWLDQVEGIRANVENFPIDVITCCSLRIRHKLGQKAFKITEQI
ESLTRQLSLISWTDDPVPLGRVGSMNASTSASSSDDFPSREKTFTQALKALEPNQQF
HMVALCGMGGVGKTRMMQRLKKAAEEKKLFNYIVRAVIGEKTDPFAIQEAIADYL
GIQLNEKTKPARADKLREWFKKNSDGGKTKFLIVLDDVWQLVDLEDIGLSPFPNQG
VDFKVLLTSRDSQVCTMMGVEANSIINVGLLTEAEAQSLFQQFVETSEPELQKIGED
IVRKCCGLPIAIKTMACTLRNKRKDAWKDALSRIEHYDIHNVAPKVFETSYHNLQE
EETKSTFLMCGLFPEDFDIPTEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQT
NLLIESDDVGCVKMHDLVRAFVLGMFSEVEHASIVNHGNMPEWTENDITDSCKRIS
LTCKSMSKFPGDFKFPNLMILKLMHGDKSLRFPQDFYEGMEKLHVISYDKMKYPLL
PLAPRCSTNIRVLHLTKCSLKMFDCSCIGNLSNLEVLSFANSRIEWLPSTVRNLKKLR
LLDLRFCDGLRIEQGVLKSLVKLEEFYIGNASGFIDDNCNEMAERSDNLSALEFAFF
NNKAEVKNMSFENLERFKISVGRSFDGNINMSSHSYENMLQLVTNKGDVLDSKLN
GLFLKTKVLFLSVHGMNDLEDVEVKSTHPTQSSSFCNLKVLIISKCVELRYLFKLNL
ANTLSRLEHLEVCECENMEELIHTGICGEETITFPKLKFLSLSQLPKLSSLCHNVNIIG
LPHLVDLILKGIPGFTVIYPQNKLRTSSLLKEEVVIPKLETLQIDDMENLEEIWPCELS
GGEKVKLREIKVSSCDKLVNLFPRNPMSLLHHLEELKVKNCGSIESLFNIDLDCVGA
IGEEDNKSLLRSINMENLGKLREVWRIKGADNSHLINGFQAVESIKIEKCKRFSNIFT
PITANFYLVALLEIQIEGCGGNHESEEQIEILSEKETLQEVTDTNISNDVVLFPSCLMH
SFHNLHKLKLERVKGVEVVFEIESESPTSRELVTTHHNQQHPIILPNLQELDLSFMD
NMSHVWKCSNWNKFFTLPKQQSESPFHNLTTIHMFSCRSIKYLFSPLMAELLSNLK
DIWISGCNGIKEVVSKRDDEDEEMTTFTSTHTTTILFPHLDSLTLRLLENLKCIGGGG
AKDEGSNEISFNNTTATTAVLDQFELSEAGGVSWSLCQYAREIEISKCNVLSSVIPCY

AAGQMQKLQVLRVTGCDGMKEVFETQLGTSSNKNRKGGGDEGNGGIPRVNNNVI MLPNLKTLKIYMCGGLEHIFTFSALESLTQLQELKIVGCYGMKVIVKKEEDEYGEQ QTTTTTTTKGASSSSSSSSSKKVVVFPRLKSIELFNLPELVGFFLGMNEFRLPSLEEVT IKYCSKMMVFAAGGSTAPQLKYIHTRLGKHTLDQESGLNFHQTSFQSLYGDTSGPA TSEGTTWSFHNLIELDMELNYDVKKIIPSSELLQLQKLEKIHVSSCYWVEEVFETAL EAAGRNGNSGIGFDESSQTTTTTTLFNLRNLREMKLHFLRGLRYIWKSNQWTAFEF PNLTRVHISRCRRLEHVFTSSMVGSLLQLQELDISWCNHMEEVIVKDADVSVEEDK ERESDGKTNKEILVLPRLKSLKLKCLPCLKGFSLGKEDFSFPLLDTLEIYKCPAITTFT KGNSATPQLKEIETRFGSFYAGEDINSSIIKRSNNRSSNKTLINVK.ILK

RG2T polynucleotide sequence (SEQ ID NO:126)

GGAAGACGACAATGGTGCAACGGTTGAAGAAGGTTGTGAAAGATAAGAAG ATGTTCCATTATATTGTCGAGGTGGTTGTAGGGGCAAACACTGACCCCAT TGCTATCCAGGATACTGTTGCAGATTACCTCAGCATAGAACTGAAAGGAA ATACGAGAGATGCAAGGGCTTATAAGCTTCGTGAATGCTTTAAGGCCCTC TCTGGTGGAGGTAAGATGAAGTTCCTAGTAATTCTTGACGATGTATGGAG CCCTGTTGATCTGGATGATATCGGTTTAAGTTCTTTGCCAAATCAAGGTG TTGACTTCAAGGTCTTGCTGACATCACGCAACAGTGATATCTGCATGATG ATGGGAGCTAGTTTAATTTTCAACCTCAATATGTTAACAGACGAGGAAGC ACATAATTTTTTCCGTCGATACGCAGAAATTTCTTATGATGCTGATCCCG AGCTTATTAAGATAGGAGAAGCTATTGTAGAGAAATGTGGTGGTTTACCC ATTGCCATCAAAACTATGGCCGTTACTCTTAGAAATAAACGCAAAGATGC ATGGAAAGATGCACTTTCTCGTTTAGAGCACCGTGACACTCATAATGTTG TGGCTGATGTTCTTAAATTGAGCTACAGCAATATCCAAGACGAGGAGACT CGGTCGATTTTTTTGCTATGTGGTTTGTTTCCTGAAGACTTTGATATTCC TACCGAAGACTTAGTGAGGTATGGATGGGGATTGAAAATATTTACCAGAG TGTATACTATGAGACATGCAAGAAAAAGGTTGGACACGTGCATTGAGCGG CTTATGCATGCCAACATGTTGATAAAAAGTGATAATGTTGGATTTGTCAA GATGCATGATCTGGTTCGTGCTTTTGTTTTGGGCATGTTATCTGAAGTCG AGCATGCATCAATTGTCAACCATGGGGATATGCCAGGGTGGTTTGAAACT GCAAATGATAAGAACAGCTTGTGCAAAAGAATTTCATTAACATGCAAAGG TATGTCTGCGATTCCTGAAGACCTCACGTTTCCAAACCTCTCGATCCTGA AATTAATGGATGGAGACGAGTCACTGAGGTTTCCTGAAGGCTTTTATGGA GAAATGGAAAACCTTCAGGTTATATCATATGATAACATGAAGCAGCCATT TCTTCCACAATCACTTCAATGCTCCAATGTTCGAGTGCTTCATCTCCATC ACTGCTCATTAATGTTTGATTGCTCTTCTATTGGAAATCTTTTGAATCTC GAGGTGCTCAGCATTGCTAATTCTGCCATTAAATTGTTACCCTCCACTAT TGGAGATCTGAAGAAGCTAAGGCTCCTGGATTTGACAAATTGTGTTGGTC TCTGTATAGCTAATGGCGTCTTTAGAAATTTGGTCAAACTTGAAGAGCTT TATATGAGAGTTGATGATCGAGATTCGTTTTTTGTGAAAGCTGATGACAG CAAGACCATTACCT

RG2T deduced polypeptide sequence (SEQ ID NO:127)

KTTMVQRLKKVVKDKKMFHYIVEVVVGANTDPIAIQDTVADYLSIELKGNTRDAR
AYKLRECFKALSGGGKMKFLVILDDVWSPVDLDDIGLSSLPNQGVDFKVLLTSRNS
DICMMMGASLIFNLNMLTDEEAHNFFRRYAEISYDADPELIKIGEAIVEKCGGLPIAI
KTMAVTLRNKRKDAWKDALSRLEHRDTHNVVADVLKLSYSNIQDEETRSIFLLCG
LFPEDFDIPTEDLVRYGWGLKIFTRVYTMRHARKRLDTCIERLMHANMLIKSDNVG
FVKMHDLVRAFVLGMLSEVEHASIVNHGDMPGWFETANDKNSLCKRISLTCKGMS
AIPEDLTFPNLSILKLMDGDESLRFPEGFYGEMENLQVISYDNMKQPFLPQSLQCSN
VRVLHLHHCSLMFDCSSIGNLLNLEVLSIANSAIKLLPSTIGDLKKLRLLDLTNCVGL
CIANGVFRNLVKLEELYMRVDDRDSFFVKADDSKTIT

RG2U polynucleotide sequence (SEQ ID NO:128)

GCCTTGTGTGGGATGGGTGGAGTGGGAAAGACCACTGTGATGAAGAAGCT
GAAGGAGGTTGTGGTAGGAAAGAAACTGTTTAATCATTATGTTGAGGCGG
TTATAGGGGAAAAGACAGACCCCATTGCTATTCAACAAGCTGTTGCCGAG
TACCTTGGTATAAGTCTAACCGAAACCACTAAACCAGCAAGAACTGATAA
GCTCCGTACATGGTTTGCAAACAACTCAAATGGAGGAAAGAAGAAGTTCC
TGGTAATACTAGACGATGTATGGCAACCAGTTGATTTGGAAGATATTGGT
TTAAGTCGTTTTCCAAATCAAGATGTTGACTTCAAGGTCTTGATTACATC
ACGGGACCAATCAGTTTGCACTGAGATGGGAGTTAAAGCTGATTTAGTTC
TCAAGGTGAGTGTCCTGGAGGAAGCGGAAGCACACAGTTTGTTCCTCCAA
TTTTTAGAACCTTCTGATGATGTCGATCCTGAGCTCAATAAAATCGGAGA
AGAAATTGTAAAGAAGTGTTGCAGACTACCCATTGCTATCAAAACCATGG
CCTGAACTCTTAGAAGTAAAAGTAAGGATACATGGAAGAATGCCCTTTCT
CGTTTACAACACCATGACATTAACACAATTGCGTCTACTGTTTTCCAAAC
TAGCTATGACAATCTCGAAGACGAGGTGACTAAAGCTACTTTTTTGCTTT
GTGGTTTATTTCCGGAGGACTTCAATATTCCTACCGAGGACCTATTGAGG
TATGGATGGGGATTGAAGTTATTCAAGGAAGTAGATACTATACGAGAAGC
AAGATCCAAGTTGAAAGCCTGCATTGAGCGGCTCATGCATACCAATTTGT
TGATCGAAGGTGATGATGTTAGGTACGTTAAGATGCATGATCTGGTGCGT
GCTTTTGTTTTGGATATGTTTTCTAAAGCCGAGCATGCATCTATTGTCAA
CCATGGTAGTAGTAAGCCAAGGTGGCCTGAAACTGAAAGTGATGTGAGCT
CCTCTTGCAAAAGAATTTCATTAACATGCAAGGGTNTG

RG2U deduced polypeptide sequence (SEQ ID NO:129)

ALCGMGGVGKTTVMKKLKEVVVGKKLFNHYVEAVIGEKTDPIAIQQAVAEYLGIS
LTETTKPARTDKLRTWFANNSNGGKKKFLVILDDVWQPVDLEDIGLSRFPNQDVD
FKVLITSRDQSVCTEMGVKADLVLKVSVLEEAEAHSLFLQFLEPSDDVDPELNKIGE
EIVKKCCRLPIAIKTMA.TLRSKSKDTWKNALSRLQHHDINTIASTVFQTSYDNLEDE
VTKATFLLCGLFPEDFNIPTEDLLRYGWGLKLFKEVDTIREARSKLKACIERLMHTN

LLIEGDDVRYVKMHDLVRAFVLDMFSKAEHASIVNHGSSKPRWPETESDVSSSCKR-ISLTCKG?

RG2V polynucleotide sequence (SEQ ID NO:130)

CTGTGGAAGACACGAATGATSAAGAAGCTGAAGGAGGTCGTGGAACAAAA
GAAAATGTTCAATATTATTGTTCAAGTGGTCATAGGAGAGAAGACAAACC
CTATTGCTATTCAGCAAGCTGTAGCAGATTACCTCTCTATTGAGCTGAAA
GAAAACACTAAAGAAGCAAGAGCTGATAAGCTTCGTNAATGGTTCGAGGA
CGATGGAGGAAAGAATAAGTTCCTTGTAATACTTGATGATGTATGGCAGT
TTGTCGATCTTGAAGATATTGGTTTAAGTCCTCTGCCAAATAAAGGTGTC
AACTTCAAGGTCTTGTTGACGTTAAGAGATTCACATGTTTGCACTCTGAT
GGGAGCTGAAGCCAATTCAATTCTCAATATAAAAGTTTTAAAAGATGTTN
AAGGACAAAGTTTGTTCCGCCAGTTTGCTAAAAATGCAGGTGATGATGAC
CTGGATCCTGCTTTCAATGGGATAGCAGATAGTATTGCAAGTAGATGTCA
AGGTTTGCCCATTGCCATCAAAACCATTGCCTTAAGTCTTAAAGGTAGAA
GCAAGCCTGCGTGGGACCATGCGCTTTCTCGTTTGGAGAACCATAAGATT
GGTAGTGAAGAAGTTGTGCGTGAAGTTTTTAAAATTAGCTATGACAATCT
CCAAGATGAGGTTACTAAATCTATTTTTWTACTTTGTGCTTTATTTCCTG
AAGATTTTGATATTCCTATTGAGGAGTTGGTGAGGTATGGGTGGGGCTTG
AAATTATTTATAGAAGCAAAAACTATAAGAGAAGCAAGAAACAGGCTCAA
CACCTGCACTGAGCGGCTTAGGGAGACAAATTTGTTATTTGGAAGTGATG
ACATTGGATGCGTCAAGATGCACGATGTGGTGCGTGATTTTGTTTGGTAT
ATATTCTCAGAAGTCCAGCACGCTTCAATTGTCAACCATGGTAATGTGTC
AGAGTGGCTAGAGGAAAATCATAGCATCTACTCTTGTAAAAGAATTTCAT
TAACATGCAAGGGTATGTCTGAGTTTCCCAAAGACCTCAAATTTCCAAAC
CTTTCAATTTTGAAACTTATGCATGGAGATAAGTCGNTGAGCTTTCCTGA-
AGACTTTTATGGAAAGATGGAAAAGGTTCAGGTAATATCATATGATAAAT
TGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCCACTAACGTTCGA
GTGCTTCATCTCCATTATTGTTCATTAAGGATGTTTGATTGCTCTTCAAT
TGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTAATTCTAACATTG
AATGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGCTACTAGAT
TTGACAAATTGTAAAGGTCTTCGTATAGATAATGGTGTCTTAAAAAATTT
GGTCAAACTTGAAGAGCTTTATATGGGTGTTAATGTCCGTATGGACCAGG
CCGT

RG2V deduced polypeptide sequence (SEQ ID NO:131)

LWKTRM?KKLKEVVEQKKMFNIIVQVVIGEKTNPIAIQQAVADYLSIELKENTKEAR
ADKLR?WFEDDGGKNKFLVILDDVWQFVDLEDIGLSPLPNKGVNFKVLLTLRDSH
VCTLMGAEANSILNIKVLKDV?GQSLFRQFAKNAGDDDLDPAFNGIADSIASRCQGL
PIAIKTIALSLKGRSKPAWDHALSRLENHKIGSEEVVREVFKISYDNLQDEVTKSIF?L
CALFPEDFDIPIEELVRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIG

CVKMHDVVRDFVWYIFSEVQHASIVNHGNVSEWLEENHSIYSCKRISLTCKGMSEF
PKDLKFPNLSILKLMHGDKS?SFPEDFYGKMEKVQVISYDKLMYPLLPSSLECSTNV
RVLHLHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKLRLLDLTNCKG
LRIDNGVLKNLVKLEELYMGVNVRMDQAV

RG2W polynucleotide sequence (SEQ ID NO:132)

TTGGGAAAGAGACAATGATGAAGAATTGAAAGAGGTTGTGGTTGAAAAGA
AAATGTTTAATCATTATGTGGAGGCGGTTATAGGGGAGAAGACGGACCCC
ATTGCTATTCAGCAAGCCGTTGCAGAGTACCTTGGTATAATTCTAACAGA
AACCACTAAGGCAGCAAGAACCGATAAGCTACGTGCATGGCTTTCTGACA
ATTCAGATGGAGGAAGAAAGAAGTTCCTAGTAATACTAGACGATGTATGG
CATCCGGTTGATATGGAAGATATTGGTTTAAGTCGTTTCCCAAATCAAGG
TGTCGACTTCAAGGTCTTGATTACATCACGGGACCAAGCTGTTTGCACTG
AGATGGGAGTTAAAGCTGATTCAGTTATCAAGGTGAGTGTCCTAGAGGAA
GCTGAAGCACAAAGCTTATTCTGCCAACTTTGGGAACCTTCTGATGATGT
CGATCCTGAGCTCCATCAGATTGGAGAAGAAATTGTAAGGAAGTGTTGTG
GTTTACCCATTGCAATAAAAACCATGGCCTGCACTCTTAGAAGTAAAAGC
AAGGATACATGGAAGAATGCACTTTCTCGTTTACAACACCATGACATTAA
CACAGTCGCGCCTACTGTTTTTCAAACCAGCTATGACAATCTCCAAGATG
AGGTGACTGGAGATACTTTTTTGCTATGTGGTTTGTTTCCGGAGGACTTC
GATATTCCTACTGAAGACTTATTGAAGTATGGATGGGGCTTAAAATTATT
CAAGGGAGTGGATTCTGTAAGAGAAGCAAGATACCAGTTGAACGCCTGCA
TTGAGCGGCTCGTGCATACCAATTTGTTGATTGAAAGTGATGTTGTTGGG
TGCGTCAAGTTGCACGATCTGGTGCGTGCTTTTATTTTGGATATGTTTTG
TAAAGCGGAGCATGCTTCGATTGTCAACCATGGTAGTAGTAAGCCTGGGT
GGCCTGAAACTGAAAATGATGTGATCAGGACCTCCTGCAAAAGAATCTCA
TTAACATGCAAGGGTATGATTGAGTTTTCTAGTGACCTCAAGTTTCCAAA
TGTCTTGATTTTAAAACTTATGCATGGAGATAAGTCGCTAAGGTTT

RG2W deduced polypeptide sequence (SEQ ID NO:133)

WERDNDEELKEVVVEKKMFNHYVEAVIGEKTDPIAIQQAVAEYLGIILTETTKAAR
TDKLRAWLSDNSDGGRKKFLVILDDVWHPVDMEDIGLSRFPNQGVDFKVLITSRD
QAVCTEMGVKADSVIKVSVLEEAEAQSLFCQLWEPSDDVDPELHQIGEEIVRKCCG
LPIAIKTMACTLRSKSKDTWKNALSRLQHHDINTVAPTVFQTSYDNLQDEVTGDTF
LLCGLFPEDFDIPTEDLLKYGWGLKLFKGVDSVREARYQLNACIERLVHTNLLIESD
VVGCVKLHDLVRAFILDMFCKAEHASIVNHGSSKPGWPETENDVIRTSCKRISLTCK
GMIEFSSDLKFPNVLILKLMHGDKSLRF

RG5 polynucleotide sequence (SEQ ID NO:134)

GGGGGGGTGGGGAAGNCGACTCTAGCCCAGAAGNTCTATAATGACCATAA
AATAAAAGGAAGCTTTAGTAAACAAGCATGGATCTGTGTTTCTCAACAAT

ATTCTGATATTTCAGTTTTGAAAGAAGTCCTTCGGAACATCGGTGTTGAT
TATAAGCATGATGAAACTGTTGGAGAACTTAGCAGAAGGCTTGCAATAGC
TGTCGAAAATGCAAGTTTCTTTCTTGTGTTGGATGATATTTGGCAACATG
AGGTGTGGACTAATTTACTCAGAGCCCCATTAAACACTGCAGCTACAGGA
ATAATTCTAGTAACAACTCGTAATGATACAGTTGCACGAGCAATTGGGGT
GGAAGATATTCATCGAGTAGAATTGATGTCAGATGAAGTAGGATGGAAAT
TGCTTTTGAAGAGTATGAACATTAGCAAAGAAAGTGAAGTAGAAAACCTA
CGAGTTTTAGGGGTTGACATTGTTCGTTTGTGTGGTGGCCTCCCCCTAGC
CTT

RG5 deduced polypeptide sequence (SEQ ID NO:135)

GGVGKTTLAQK?YNDHKIKGSFSKQAWICVSQQYSDISVLKEVLRNIGVDYKHDET
VGELSRRLAIAVENASFFLVLDDIWQHEVWTNLLRAPLNTAATGIILVTTRNDTVA
RAIGVEDIHRVELMSDEVGWKLLLKSMNISKESEVENLRVLGVDIVRLCGGLPLAL

RG7 polynucleotide sequence (SEQ ID NO:136)

GGTGGGGTTGGGAAGACAACGGGCACAAGGAGGCGACTGCCAATACTTCC
GACTTTTATTCATAGAGATGACGAGTCTTATTTTCCTACTACTATAGGGA
GGATATTTGGTTGCGCGAGACGATTCATTGCGCGAAGGGATTCTATCCTT
CTTTTTTTCCGCGAAGACTTCGTTCCGGAGGACGGGCTATATTCCCTTTA
ATATTAGTCTAGCCCAGTCTAGGCCAACCATATGGCGATGCGGTAGACCT
CCCAGAGATAGATACTTGATCTTAGAGGATTCACACGTTCAATGGTGGAA
ACTTAAGGAACCGGCTAAGAGTGACTAAACGGAAAAACCCTATTCATTCC
ATAGCCTCATCCGGTCGAGGCATTAAACAATCCATCCCAATCCTCTTTCC
TTTGGTCTACTCTAATGATGTGCCCGTTCGTTGGTGGAATATCTCTTTAT
ACCGACGATTTATATGGGGATTGCCACTAGCGTTG

The above examples are provided to illustrate the invention but not to limit its scope. Other variants of the invention will be readily apparent to one of ordinary skill in the art and are encompassed by the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid construct comprising an RG polynucleotide which encodes an RG polypeptide having at least 60% sequence identity to an RG polypeptide from an RG family selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide.

2. The nucleic acid construct of claim 1, wherein the RG polynucleotide encodes an RG polypeptide comprising an leucine rich region (LRR).

3. The nucleic acid construct of claim 1, wherein the RG polynucleotide encodes an RG polypeptide comprising a nucleotide binding site (NBS).

4. The nucleic acid construct of claim 1, wherein the polynucleotide is a full length gene.

5. The nucleic acid construct of claim 1, wherein the further encodes a fusion protein.

6. The nucleic acid construct of claim 1, wherein the RG1 polypeptide is encoded by an RG1 polynucleotide sequence.

7. The nucleic acid construct of claim 6, wherein the RG1 polypeptide is encoded by a polynucleotide sequence selected from the group consisting of SEQ ID NO:1 (RG1A), SEQ ID NO:2 (RG1B), SEQ ID NO: 3 (RG1C), SEQ ID NO:4 (RG1D), SEQ ID NO:5 (RG1E), SEQ ID NO:6 (RG1F), SEQ ID NO:7 (RG1G), SEQ ID NO:8 (RG1H), SEQ ID NO:9 (RG1I), and SEQ ID NO:10 (RG1J).

8. The nucleic acid construct of claim 1, wherein the RG2 polypeptide is encoded by an RG2 polynucleotide sequence.

9. The nucleic acid construct of claim 8, wherein the RG2 polypeptide is encoded by a polynucleotide sequence selected from the group consisting of: SEQ ID NO:21 (RG2A);

SEQ ID NO:23 (RG2B); SEQ ID NO:25 (RG2C); SEQ ID NO:27 (RG2D); SEQ ID NO:29 (RG2E); SEQ ID NO:31 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:35 (RG2H); SEQ ID NO:37 (RG2I); SEQ ID NO:39 (RG2J); SEQ ID NO:41 (RG2K); SEQ ID NO:43 (RG2L); SEQ ID NO:45 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).

10. The nucleic acid construct of claim 1, wherein the RG3 polypeptide is encoded by an RG3 polynucleotide sequence.

11. The nucleic acid construct of claim 10, wherein the RG3 polypeptide is encoded by a polynucleotide sequence as set forth in SEQ ID NO:68.

12. The nucleic acid construct of claim 1, wherein the RG4 polypeptide is encoded by an RG4 polynucleotide sequence.

13. The nucleic acid construct of claim 12, wherein the RG4 polypeptide is encoded by a polynucleotide sequence as set forth in SEQ ID NO:69.

14. The nucleic acid construct of claim 1, wherein the RG5 polypeptide is encoded by an RG5 polynucleotide sequence.

15. The nucleic acid construct of claim 14, wherein the RG5 polypeptide is encoded by a polynucleotide sequence as set forth in SEQ ID NO:134.

16. The nucleic acid construct of claim 1, wherein the RG7 polypeptide is encoded by an RG7 polynucleotide sequence.

17. The nucleic acid construct of claim 16, wherein the RG7 polypeptide is encoded by a polynucleotide sequence as set forth in SEQ ID NO:136.

18. The nucleic acid construct of claim 1, further comprising a promoter operably linked to the RG polynucleotide.

19. The nucleic acid construct of claim 18, wherein the promoter is a plant promoter.

20. The nucleic acid construct of of claim 19, wherein the plant promoter is a disease resistance promoter.

21. The nucleic acid construct of claim 19, wherein the plant promoter is a lettuce promoter.

22. The nucleic acid construct of claim 18, wherein the promoter is a constitutive promoter.

23. The nucleic acid construct of claim 18, wherein the promoter is an inducible promoter.

24. The nucleic acid construct of claim 18, wherein the promoter is a tissue-specific promoter.

25. A nucleic acid construct comprising a promoter sequence from an RG gene linked to a heterologous polynucleotide.

26. A transgenic plant comprising a recombinant expression cassette comprising a promoter operably linked to an RG polynucleotide.

27. The transgenic plant of claim 26, wherein the plant promoter is a plant promoter.

28. The transgenic plant of claim 26, wherein the plant promoter is a viral promoter.

29. The transgenic plant of claim 26, wherein the plant promoter is a heterologous promoter.

30. The transgenic plant of claim 26, wherein the plant is lettuce.

31. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:1 (RG1A), SEQ ID NO:2 (RG1B), SEQ ID NO: 3 (RG1C), SEQ ID NO:4 (RG1D), SEQ ID NO:5 (RG1E), SEQ ID NO:6 (RG1F), SEQ ID NO:7 (RG1G), SEQ ID NO:8 (RG1H), SEQ ID NO:9 (RG1I), and SEQ ID NO:10 (RG1J).

32. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:21 (RG2A); SEQ ID NO:23 (RG2B); SEQ ID NO:25 (RG2C); SEQ ID NO:27 (RG2D); SEQ ID NO:29 (RG2E); SEQ ID NO:31 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:35 (RG2H); SEQ ID NO:37 (RG2I); SEQ ID NO:39 (RG2J); SEQ ID NO:41 (RG2K); SEQ ID NO:43 (RG2L); SEQ ID NO:45 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).

33. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:68 (RG3) and SEQ ID NO:69 (RG4).

34. The transgenic plant of claim 26, wherein the RG polynucleotide comprises a sequence as set forth in SEQ ID NO:134 (RG5).

35. The transgenic plant of claim 26, wherein the RG polynucleotide comprises a sequence as set forth in SEQ ID NO:136 (RG7).

36. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG1 polypeptide selected from the group consisting of SEQ ID NO:11 (RG1A), SEQ ID NO:12 (RG1B), SEQ ID NO: 13 (RG1C), SEQ ID NO:14 (RG1D), SEQ ID NO:15 (RG1E), SEQ ID NO:16 (RG1F), SEQ ID NO:17 (RG1G), SEQ ID NO:18 (RG1H), SEQ ID NO:19 (RG1I), and SEQ ID NO:20 (RG1J).

37. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG2 polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W).

38. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG3 polypeptide with a sequence as set forth by SEQ ID NO:138.

39. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG4 polypeptide with a sequence as set forth by SEQ ID NO:139.

40. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG5 polypeptide with a sequence as set forth by SEQ ID NO:135.

41. A method of enhancing disease resistance in a plant, the method comprising introducing into the plant a recombinant expression cassette comprising a promoter functional in the plant and operably linked to an RG polynucleotide sequence.

42. The method of claim 41, wherein the plant is a lettuce plant.

43. The method of claim 41, wherein the RG polynucleotide encodes an RG polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W).

44. The method of claim 41, wherein the RG polynucleotide encodes an RG polypeptide selected from the group consisting of SEQ ID NO:138 (RG3); SEQ ID NO:139 (RG4); and SEQ ID NO:135 (RG5).

45. The method of claim 41, wherein the promoter is a tissue-specific promoter or a plant disease resistance promoter.

46. The method of claim 41, wherein the promoter is a constitutive promoter or an inducible promoter.

47. A method of detecting RG resistance genes in a nucleic acid sample, the method comprising:

contacting the nucleic acid sample with an RG polynucleotide to form a hybridization complex; and,

wherein the formation of the hybridization complex is used to detect the RG resistance gene in the nucleic acid sample.

48. The method of claim 47, wherein the RG polynucleotide is an RG1 polynucleotide.

49. The method of claim 47, wherein the RG polynucleotide is an RG2 polynucleotide.

50. The method of claim 47, wherein the RG polynucleotide is an RG3 polynucleotide, an RG4 polynucleotide, an RG5 polynucleotide or an RG7 polynucleotide.

51. The method of claim 47, wherein the RG resistance gene is amplified prior to the step of contacting the nucleic acid sample with the RG polynucleotide.

52. The method of claim 51, where the RG resistance gene is amplified by the polymerase chain reaction.

53. The method of claim 47, wherein the RG polynucleotide is labeled.

54. An RG polypeptide having at least 60% sequence identity to a polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/00615

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :435/6, 91.2, 418, 419; 530/350; 536/23.1, 23.6, 24.1; 800/205
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 91.2, 418, 419; 530/350; 536/23.1, 23.6, 24.1; 800/205

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PARAN et al. Development of Reliable PCR-Based Markers Linked to Downy Mildew Resistance Genes in Lettuce. Theor. Appl. Genet. 1993. Vol. 85, No. 8, pages 985-993, see entire article.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54
Y	KESSELI et al. Analysis of a Detailed Genetic Linkage Map of Lactuca sativa (Lettuce) Constructed From RFLP and RAPD Markers. Genetics. April 1994. Vol. 136, No. 4, pages 1435-1446, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54
Y	MICHELMORE, RW. Isolation of Disease Resistance Genes from Crop Plants. Current Opinion in Biotechnology. 1995. Vol. 6, No. 2, pages 145-152, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:
- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed
- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 MARCH 1998	13 APR 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
PHUONG BUI
Telephone No. (703) 308-0196

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PARAN et al. Recent Amplification of Triose Phosphate Isomerase Related Sequences in Lettuce. Genome. 1992. Vol. 35, No. 4, pages 627-635, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54
Y	PARAN et al. Identification of Restriction Fragment Length Polymorphism and Random Amplified Polymorphic DNA markers linked to Downy Mildew Resistance Genes in Lettuce, Using Near-Isogenic Lines. Genome. 1991. Vol. 34, No. 6, pages 1021-1027, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/00615

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 7, 9, 11, 13, 15, 17, 31-40, 43-44
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

these claims are drawn to numerous sequences identified by SEQ ID NOs. However, since no computer readable form was submitted, no meaningful search could be carried out.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/00615

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

A01H 1/00; C07H 21/04; C07K 14/00; C12N 5/04, 5/10; C12P 19/34; C12Q 1/68